```
            GenCore version 5.1.9
        Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on:    July 11, 2006, 07:21:26 ; Search time 107.065 Seconds
                                (without alignments)
                          4561.330 Million cell updates/sec

Title:          US-09-973-994-131
Perfect score:  261
Sequence:       1 aggtgaccgtacagcatta.........gggtttttccacggtcacct  261

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters:    2807332

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    Issued_Patents_NA:*
        1:  /EMC_Celerra_SIDS3/ptodata/2/ina/1_COMB.seq:*
        2:  /EMC_Celerra_SIDS3/ptodata/2/ina/5_COMB.seq:*
        3:  /EMC_Celerra_SIDS3/ptodata/2/ina/6A_COMB.seq:*
        4:  /EMC_Celerra_SIDS3/ptodata/2/ina/6B_COMB.seq:*
        5:  /EMC_Celerra_SIDS3/ptodata/2/ina/7_COMB.seq:*
        6:  /EMC_Celerra_SIDS3/ptodata/2/ina/H_COMB.seq:*
        7:  /EMC_Celerra_SIDS3/ptodata/2/ina/PCTUS_COMB.seq:*
        8:  /EMC_Celerra_SIDS3/ptodata/2/ina/PP_COMB.seq:*
        9:  /EMC_Celerra_SIDS3/ptodata/2/ina/RE_COMB.seq:*
       10:  /EMC_Celerra_SIDS3/ptodata/2/ina/backfiles1.seq:*

                    SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

Result No.		Score	Query Match %	Query Match Length	DB	ID	Description
c	1	37.6	14.4	144362	3	US-09-949-016-16066	Sequence 16066, A
	2	36.8	14.1	357	3	US-09-640-211A-1581	Sequence 1581, Ap
	3	35.8	13.7	264206	3	US-09-949-016-12731	Sequence 12731, A
	4	35.8	13.7	264304	3	US-09-949-016-13249	Sequence 13249, A
c	5	35.6	13.6	117391	3	US-09-949-016-13945	Sequence 13945, A
c	6	34.8	13.3	139049	3	US-09-949-016-17030	Sequence 17030, A
	7	34.8	13.3	193555	3	US-09-949-016-15553	Sequence 15553, A
	8	34.8	13.3	193555	3	US-09-949-016-15554	Sequence 15554, A
	9	34.8	13.3	193555	3	US-09-949-016-15555	Sequence 15555, A
	10	34.6	13.3	200663	3	US-09-949-016-12569	Sequence 12569, A
	11	34.6	13.3	203093	3	US-09-949-016-14445	Sequence 14445, A
c	12	33.6	12.9	601	3	US-09-949-016-53593	Sequence 53593, A
	13	33.6	12.9	57392	3	US-09-949-016-12070	Sequence 12070, A
	14	33.6	12.9	57402	3	US-09-949-016-13293	Sequence 13293, A
c	15	33.4	12.8	786	3	US-09-134-000C-905	Sequence 905, App
c	16	33.2	12.7	43360	3	US-09-453-702B-206	Sequence 206, App
c	17	33.2	12.7	43360	3	US-10-114-170-206	Sequence 206, App
c	18	33.2	12.7	45325	3	US-09-453-702B-261	Sequence 261, App
c	19	33.2	12.7	45325	3	US-10-114-170-261	Sequence 261, App
c	20	33.2	12.7	156324	3	US-09-949-016-13749	Sequence 13749, A
c	21	33	12.6	237863	3	US-09-949-016-13404	Sequence 13404, A
c	22	32.8	12.6	150	5	US-09-974-300-7683	Sequence 7683, Ap
c	23	32.8	12.6	56374	3	US-09-949-002-645	Sequence 645, App
c	24	32.8	12.6	56375	3	US-09-949-002-774	Sequence 774, App
c	25	32.8	12.6	202001	3	US-09-734-674-3	Sequence 3, Appli
c	26	32.6	12.6	202001	3	US-10-274-990-3	Sequence 3, Appli
c	27	32.6	12.5	56976	3	US-09-949-016-17486	Sequence 17486, A
	28	32.6	12.5	98567	3	US-09-949-016-11750	Sequence 11750, A
	29	32.6	12.5	100567	3	US-09-949-016-16934	Sequence 16934, A
c	30	32.4	12.4	101674	3	US-09-949-016-12033	Sequence 12033, A
c	31	32.4	12.4	601	3	US-09-949-016-113225	Sequence 113225,
	32	32.4	12.4	285986	3	US-09-949-016-12287	Sequence 12287, A
	33	32.4	12.4	288031	3	US-09-949-016-14864	Sequence 14864, A
c	34	32.4	12.4	340380	3	US-09-949-016-14179	Sequence 14179, A
	35	32.2	12.3	436	3	US-09-513-999C-30681	Sequence 30681, A
	36	32.2	12.3	1709	3	US-09-270-767-14410	Sequence 14410, A
	37	32	12.3	39982	3	US-09-820-924-3	Sequence 3, Appli
	38	32	12.3	39982	3	US-10-369-626-3	Sequence 3, Appli
	39	32	12.3	51043	3	US-09-949-016-12739	Sequence 12739, A
	40	32	12.3	51046	3	US-09-949-016-13946	Sequence 13946, A
	41	32	12.3	144158	3	US-09-949-016-11755	Sequence 11755, A
	42	32	12.3	144158	3	US-09-949-016-12936	Sequence 12936, A
c	43	31.8	12.2	1664976	3	US-08-916-421B-1	Sequence 1, Appli
c	44	31.8	12.2	1664976	3	US-09-692-570-1	Sequence 1, Appli
c	45	31.6	12.1	366	3	US-08-956-171B-2607	Sequence 2607, Ap

```
                    ALIGNMENTS

RESULT 1
US-09-949-016-16066/c
; Sequence 16066, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;   APPLICANT: VENTER, J. Craig et al.
;   TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;   TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;   FILE REFERENCE: CL001307
;   CURRENT APPLICATION NUMBER: US/09/949,016
;   CURRENT FILING DATE:  2000-04-14
;   PRIOR APPLICATION NUMBER: 60/241,755
;   PRIOR FILING DATE:  2000-10-20
;   PRIOR APPLICATION NUMBER: 60/237,768
;   PRIOR FILING DATE:  2000-10-03
;   PRIOR APPLICATION NUMBER: 60/231,498
;   PRIOR FILING DATE:  2000-09-08
;   NUMBER OF SEQ ID NOS: 207012
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16066
;   LENGTH: 144362
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(144362)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16066

Query Match            14.4%;  Score 37.6;  DB 3;  Length 144362;
Best Local Similarity  53.4%;  Pred. No. 1.1;
Matches  79;  Conservative  0;  Mismatches  69;  Indels  0;  Gaps  0;

Qy          92 AAAGTTGCTCGCTGATTTGAGGCACGCGAAACGAGATAAGTTGTAAAATTTGCTCGCT 151
               |||||||  |||||||||| || ||  |   |   | |||||| |||  |||||||||
Db      105133 ACAGGTGAATTGATTATGTAAACCAACTAATATGGACGAGAAATATAATTAATATTCT 105074

Qy         152 GATTTTTGCTGAATATTTCTCACTATAAAAGCATTTCCAGAATAAGAAGGAGCT 211
               |||   ||  |||| | || |||  ||||||| ||  |||| |||||| |||||
Db      105073 ATTGTTAGGTTAGCAAATTACTTACAATAAAATCAGTTATCAATCAATCAGTTATCAATCATTATAAGGGCTT 105014

Qy         212 TTCGAACTGGTTTTCCCCAAGAGTTGTA 239
               ||||||||||| |||||| |||||||
Db      105013 TTCAAAAAGCATGTCAAAAGACAGTA 104986
```

Db 14035 AAACTAATCCAGACACAAGGTGTCAGTTTGGAAATAATGAAGCATATTTCGTTTTAA 13976
           |||| || |||| || || ||| ||| ||| || |||| |||| |
Qy   165 ATATTTCTCCACTATAAAAAGCATTTTCCA 195
           || | |||||||||| ||| |||
Db 13975 AAATCTAAAACATTTTAAAGACAAAAATTCA 13945

RESULT 4
US-09-949-016-13249/c
; Sequence 13249, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: PatentSEQ for Windows Version 4.0
; SEQ ID NO 13249
; LENGTH: 264304
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13249

Query Match   13.7%; Score 35.8; DB 3; Length 264304;
Best Local Similarity   52.3%; Pred. No. 4.3;
Matches   79; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

Qy   45 AAGTTTTCCGATTCGCTGTGAGGCACGGAAAACGAGATAAGTTGTAAAAGTTTGCTCGC 104
           ||| |||| ||| || || |||| ||| |||| ||||| ||
Db 14095 AAATATTTCAAAATCAATGGACTTCAAGTAATCCAAGAGAATTTGGAAAACATTTCTCTA 14036
Qy   105 TGATTTGAGGCACGGAAAACGAGATAAGTTGTAAAATTTGCTCGCTGATTTTTTGCTGA 164
           | |||||||||||||||||||||
Db 14035 AAACTAATCCAGACACAAGGTGTCAGTTTGGAAATAATGAAGCATATTTCGTTTTAA 13976
Qy   165 ATATTTCTCCACTATAAAAAGCATTTTCCA 195
           || | |||||||||| ||| |||
Db 13975 AAATCTAAAACATTTTAAAGACAAAAATTCA 13945

RESULT 5
US-09-949-016-13945/c
; Sequence 13945, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: PatentSEQ for Windows Version 4.0
; SEQ ID NO 13945
; LENGTH: 117391
; TYPE: DNA
; ORGANISM: Human

RESULT 2
US-09-640-211A-1581
; Sequence 1581, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: PatentSEQ for Windows Version 4.0
; SEQ ID NO 1581
; LENGTH: 357
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-1581

Query Match   14.1%; Score 36.8; DB 3; Length 357;
Best Local Similarity   52.6%; Pred. No. 0.31;
Matches   80; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

Qy   46 AGTTTTCCGATTCGCTGTGAGGCACGGAAAACGAGATAAGTTGTAAAAGTTTGCTCGCT 105
           |||||||||||||||||| || |||| | || |||| ||||| || |
Db   85 AGTTCATCTGGTTCAGTTTGGACTATACAACAGAACAAAATTTTTGAAAATGCTCTAGCT 144

Qy   106 GATTTGAGGCACGGAAAACGAGATAAGTTGTAAAATTTGCTCGCTGATTTTTTGCTGAA 165
           ||||| |||||||||||||||||||||
Db   145 GATTTTGATAAGACACCCCAGATAAATGGGAGAAGTGGCAGCCAGGCTGCCTGAAAA   204

Qy   166 TATTTCTCCACTATAAAAGCATTTTCCAGA 197
           ||| | |||||||||| |||
Db   205 ACTGCTACGGATGTTAGAAGCATTATGAAGA 236

RESULT 3
US-09-949-016-12731/c
; Sequence 12731, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: PatentSEQ for Windows Version 4.0
; SEQ ID NO 12731
; LENGTH: 264206
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12731

Query Match   13.7%; Score 35.8; DB 3; Length 264206;
Best Local Similarity   52.3%; Pred. No. 4.3;
Matches   79; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

Qy   45 AAGTTTTCCGATTCGCTGTGAGGCACGGAAAACGAGATAAGTTGTAAAAGTTTGCTCGC 104
           ||| |||| ||| || || |||| ||| |||| ||||| ||
Db 14095 AAATATTTCAAAATCAATGGACTTCAAGTAATCCAAGAGAATTTGGAAAACATTTCTCTA 14036
Qy   105 TGATTTGAGGCACGGAAAACGAGATAAGTTGTAAAATTTGCTCGCTGATTTTTTGCTGA 164

```
                    GenCore version 5.1.9
          Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on:    July 11, 2006, 07:46:06 ; Search time 53.7623 Seconds
                                    (without alignments)
                                 6311.756 Million cell updates/sec

Title:        US-09-973-994-131
Perfect score: 261
Sequence:     1 aggtgaccgtacagcattta.........gggtttccacggtcacct 261

Scoring table: IDENTITY_NUC
              Gapop 10.0 , Gapext 1.0

Searched:     844069 seqs, 650066433 residues

Total number of hits satisfying chosen parameters:     1688138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    Published Applications_NA_New:*
              1: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US09_NEW_PUB.seq:*
              2: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US06_NEW_PUB.seq:*
              3: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US07_NEW_PUB.seq:*
              4: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US08_NEW_PUB.seq:*
              5: /EMC_Celerra_SIDS3/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
              6: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US10_NEW_PUB.seq:*
              7: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq:*
              8: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq1:*
              9: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                          SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description
1	36.4	13.9	962	8	US-11-266-748A-97714	Sequence 97714, A
2	36.4	13.9	962	8	US-11-266-748A-150525	Sequence 150525,
3	36.4	13.9	1421	8	US-11-266-748A-361361	Sequence 361361,
4	36.4	13.9	1421	8	US-11-266-748A-444740	Sequence 444740,
5	36.4	13.9	1888	8	US-11-266-748A-260726	Sequence 260726,
6	36.4	13.9	1888	8	US-11-266-748A-321243	Sequence 321243,
7	35.8	13.7	988	8	US-11-266-748A-192101	Sequence 192101,
8	35.8	13.7	988	8	US-11-266-748A-226177	Sequence 226177, A
9	35.4	13.6	686	8	US-11-266-748A-55772	Sequence 55772, A
10	34.4	13.2	706	8	US-11-266-748A-166299	Sequence 166299,
11	34	13.0	489	8	US-11-266-748A-172138	Sequence 172138, A
12	34	13.0	888	8	US-11-266-748A-19927	Sequence 19927, A
13	33.8	13.0	2079	8	US-11-266-748A-180693	Sequence 180693, A
14	33.6	12.9	3350	8	US-11-266-748A-28652	Sequence 28652, A
15	32.6	12.5	1000	8	US-11-266-748A-118940	Sequence 118940,
16	32.6	12.5	1000	8	US-11-266-748A-161104	Sequence 161104,
17	32.6	12.5	1000	8	US-11-266-748A-224558	Sequence 224558,
18	32.6	12.5	1000	8	US-11-266-748A-293755	Sequence 293755,
19	32.6	12.5	1000	8	US-11-266-748A-345184	Sequence 345184,
20	32.6	12.5	1000	8	US-11-266-748A-396868	Sequence 396868,
21	32.6	12.5	1000	8	US-11-266-748A-405877	Sequence 405877,
22	32.6	12.5	1000	8	US-11-266-748A-467914	Sequence 467914,
23	32.6	12.5	1000	6	US-11-266-748A-476923	Sequence 476923,
24	32.6	12.5	2199	6	US-10-471-571A-421	Sequence 421, App

25	32.6	12.5	10884	8	US-11-266-748A-29503	Sequence 29503, A
26	32.6	12.5	10935	8	US-11-266-748A-56647	Sequence 56647, A
27	32.2	12.3	795	8	US-11-266-748A-87473	Sequence 87473, A
28	32.2	12.3	795	8	US-11-266-748A-140284	Sequence 140284,
29	32.2	12.3	2195	8	US-11-293-697-1009	Sequence 1009, Ap
30	32.2	12.3	2519	8	US-11-293-697-2370	Sequence 2370, Ap
31	32.2	12.3	3609	8	US-11-266-748A-31160	Sequence 31160, A
32	32.2	12.3	3862	8	US-11-266-748A-31694	Sequence 31694, A
33	32.2	12.3	29492	7	US-11-001-837-23	Sequence 23, Appl
34	32.2	12.3	909203	8	US-11-266-748A-29039	Sequence 29039, A
35	31.6	12.1	1608	6	US-10-471-571A-253	Sequence 253, App
36	31.6	12.1	1930	8	US-11-266-748A-26679	Sequence 26679, A
37	31.6	12.1	3732	8	US-11-293-697-826	Sequence 826, App
38	31.4	12.0	239	8	US-11-266-748A-420002	Sequence 420002,
39	31.4	12.0	381	8	US-11-266-748A-11128	Sequence 11128, A
40	31.4	12.0	462	8	US-11-266-748A-369370	Sequence 369370,
41	31.4	12.0	462	8	US-11-266-748A-457749	Sequence 457749,
42	31.4	12.0	464	8	US-11-266-748A-70104	Sequence 70104, A
43	31.4	12.0	464	8	US-11-266-748A-122915	Sequence 122915,
44	31.4	12.0	644	8	US-11-266-748A-182148	Sequence 182148,
45	31.4	12.0	644	8	US-11-266-748A-242695	Sequence 242695,

```
                                                ALIGNMENTS

RESULT 1
US-11-266-748A-97714
; Sequence 97714, Application US/11266748A
; Publication No. US200601344663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT: Johnston, Patrick
;  APPLICANT: Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION: Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE: 2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 97714
;  LENGTH: 962
;  TYPE: DNA
;  ORGANISM: Homo Sapiens
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (231)..(231)
;  OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-97714

Query Match             13.9%;  Score 36.4;  DB 8;  Length 962;
Best Local Similarity   53.5%;  Pred. No. 0.77;
Matches    76;  Conservative    0;  Mismatches   66;  Indels    0;  Gaps    0;

Qy       75 AAACGAGATAAGTTGTAAAAGTTGCTCGCTGATTTGAGGCACGGAAAACGAGATAAGTT 134
            | ||||  ||||||| ||| ||| |||  ||||   |||||  || ||||   ||||||
```

```
Db   799 ACACAGCCTCATGGTAAAATTTTCCACACTGAGTTACCACCAGCGCTTAACATCTGTC 858
Qy   135 GTAAAATTTTGCTCGCTGATTTTTGCTGAATATTTCTCTCACTATAAAAAGCATTTTCC 194
         ||||  ||| |||||  |||     ||  ||||||||||||||||||||||
Db   859 TTGCTCTCTTTACACACGAAACATGAGTGAATCCCTTGCTCACTAGTCCAAGCATTCTCC 918

Qy   195 AGAAATAAGAAGGAGCTTTCGA 216
         ||||| |||| ||| |||
Db   919 AGCAAAAAGATCCAGCCATCAA 940
```

RESULT 2
US-11-266-748A-150525/c
; Sequence 150525, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 150525
; LENGTH: 962
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (732)..(732)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-150525

Query Match   13.9%; Score 36.4; DB 8; Length 962;
Best Local Similarity   53.5%; Pred. No. 0.77;
Matches   76; Conservative   0; Mismatches 66; Indels   0; Gaps   0;

```
Qy    75 AAACGAGATAAGTTGTAAAAGTTGCTCGCTGATTTGAGGCACGGAAAACGAGATAAGTT 134
         |||||  ||||  ||||   ||| ||  ||||   |||
Db   164 ACACAGCCTCATGGTAAAATTTTCCACACTGAGTTACCACCAGCGCTTAACATCTGTC 105

Qy   135 GTAAAATTTTGCTCGCTGATTTTTGCTGAATATTTCTCTCACTATAAAAAGCATTTTCC 194
         ||||  ||| |||||  |||     ||  ||||||||||||||||||||||
Db   104 TTGCTCTCTTTACACACGAAACATGAGTGAATCCCTTGCTCACTAGTCCAAGCATTCTCC 45

Qy   195 AGAAATAAGAAGGAGCTTTCGA 216
         ||||| |||| ||| |||
Db    44 AGCAAAAAGATCCAGCCATCAA 23
```

RESULT 3
US-11-266-748A-361361/c
; Sequence 361361, Application US/11266748A
; Publication No. US20060134663A1

GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 361361
; LENGTH: 1421
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-361361

Query Match   13.9%; Score 36.4; DB 8; Length 1421;
Best Local Similarity   53.5%; Pred. No. 0.87;
Matches   76; Conservative   0; Mismatches 66; Indels   0; Gaps   0;

```
Qy    75 AAACGAGATAAGTTGTAAAAGTTGCTCGCTGATTTGAGGCACGGAAAACGAGATAAGTT 134
         |||||  ||||  ||||   ||| ||  ||||   |||
Db   164 ACACAGCCTCATGGTAAAATTTTCCACACTGAGTTACCACCAGCGCTTAACATCTGTC 105

Qy   135 GTAAAATTTTGCTCGCTGATTTTTGCTGAATATTTCTCTCACTATAAAAAGCATTTTCC 194
         ||||  ||| |||||  |||     ||  ||||||||||||||||||||||
Db   104 TTGCTCTCTTTACACACGAAACATGAGTGAATCCCTTGCTCACTAGTCCAAGCATTCTCC 45

Qy   195 AGAAATAAGAAGGAGCTTTCGA 216
         ||||| |||| ||| |||
Db    44 AGCAAAAAGATCCAGCCATCAA 23
```

RESULT 4
US-11-266-748A-444740
; Sequence 444740, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03

```
                      GenCore version 5.1.9
            Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on:     July 11, 2006, 07:46:32 ; Search time 779.324 Seconds
                                      (without alignments)
                                      4115.199 Million cell updates/sec

Title:          US-09-973-994-131
Perfect score:  261
Sequence:       1 aggtgaccgtacagcatta..........gggtttttccacggtcacct 261

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters:    37784340

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :     Published Applications NA Main:*
          1: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US07_PUBCOMB.seq:*
          2: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US08_PUBCOMB.seq:*
          3: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US09A_PUBCOMB.seq:*
          4: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US09B_PUBCOMB.seq:*
          5: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US09C_PUBCOMB.seq:*
          6: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US10A_PUBCOMB.seq:*
          7: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US10B_PUBCOMB.seq:*
          8: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US10C_PUBCOMB.seq:*
          9: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US10C_PUBCOMB.seq:*
         10: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US10D_PUBCOMB.seq:*
         11: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US10E_PUBCOMB.seq:*
         12: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US10F_PUBCOMB.seq:*
         13: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US10G_PUBCOMB.seq:*
         14: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US11A_PUBCOMB.seq:*
         15: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US11B_PUBCOMB.seq:*
         16: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US11C_PUBCOMB.seq:*
             /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US11D_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Query Match Length	DB	ID	Description
1	261	100.0	261	10	US-10-651-991-131	Sequence 131, App
2	37.4	14.3	40324	8	US-10-433-793-180	Sequence 180, App
3	37	14.2	346	4	US-09-925-065A-125478	Sequence 125478,
4	37	14.2	346	4	US-09-925-065A-125480	Sequence 125480,
5	37	14.2	346	5	US-09-925-065A-125481	Sequence 125481,
6	37	14.2	346	5	US-09-925-065A-125478	Sequence 125478,
7	37	14.2	346	5	US-09-925-065A-125480	Sequence 125480,
8	37	14.2	346	5	US-09-925-065A-125481	Sequence 125481,
9	37	14.2	558	4	US-09-925-065A-439656	Sequence 439656,
10	37	14.2	558	5	US-09-925-065A-439656	Sequence 439656,
11	36.8	14.1	357	9	US-10-856-499-1581	Sequence 1581, Ap
12	36.8	14.1	5276	8	US-10-311-455-800	Sequence 800, App
13	36.8	14.1	5276	7	US-10-221-714A-102	Sequence 102, App
14	36.8	14.1	40862	7	US-10-311-455-2046	Sequence 2046, Ap
15	36.6	14.0	346	4	US-09-925-065A-125479	Sequence 125479,
16	36.6	14.0	346	5	US-09-925-065A-125479	Sequence 125479,
17	36	13.8	37515	8	US-10-433-793-27	Sequence 27, Appl

c	18	35.8	13.7	608	4	US-09-925-065A-787534	Sequence 787534,
c	19	35.8	13.7	608	5	US-09-925-065A-787534	Sequence 787534,
	20	35.8	13.7	611	4	US-09-925-065A-73437	Sequence 73437, A
	21	35.8	13.7	611	5	US-09-925-065A-73437	Sequence 73437, A
	22	35.8	13.7	611	12	US-10-301-480-174676	Sequence 174676,
	23	35.8	13.7	611	12	US-10-301-480-788085	Sequence 788085,
	24	35.8	13.7	654	4	US-09-925-065A-778844	Sequence 778844,
	25	35.8	13.7	654	4	US-09-925-065A-778843	Sequence 778843,
	26	35.8	13.7	654	5	US-09-925-065A-778843	Sequence 778843,
	27	35.8	13.7	654	5	US-09-925-065A-778844	Sequence 778844,
	28	35.8	13.7	680	4	US-09-925-065A-785063	Sequence 785063,
	29	35.8	13.7	680	4	US-09-925-065A-785064	Sequence 785064,
	30	35.8	13.7	680	4	US-09-925-065A-785065	Sequence 785065,
	31	35.8	13.7	680	5	US-09-925-065A-785063	Sequence 785063,
	32	35.8	13.7	680	5	US-09-925-065A-785064	Sequence 785064,
	33	35.8	13.7	680	5	US-09-925-065A-785065	Sequence 785065,
	34	35.8	13.7	823	6	US-10-027-632-172361	Sequence 172361,
c	35	35.8	13.7	823	6	US-10-027-632-172361	Sequence 172361,
c	36	35.8	13.7	2325	6	US-10-027-632-103528	Sequence 103528,
	37	35.8	13.7	2325	6	US-10-027-632-103529	Sequence 103529,
	38	35.8	13.7	2325	7	US-10-027-632-103528	Sequence 103528,
	39	35.8	13.7	2325	7	US-10-027-632-103529	Sequence 103529,
c	40	35.8	13.7	561515	8	US-10-741-601-5682	Sequence 5682, Ap
c	41	35.8	13.7	561515	9	US-10-741-600-17730	Sequence 17730, A
c	42	35.8	13.7	3673778	7	US-10-312-841-2	Sequence 2, Appli
c	43	35.4	13.6	969	12	US-10-301-480-570600	Sequence 570600,
c	44	35.4	13.6	969	12	US-10-301-480-1184009	Sequence 1184009,
c	45	35.4	13.6	987	12	US-10-301-480-603635	Sequence 603635,

ALIGNMENTS

```
RESULT 1
US-10-651-991-131
; Sequence 131, Application US/10651991
; Publication No. US20050125161A1
; GENERAL INFORMATION:
;  APPLICANT: CAIRNEY, JOHN
;  APPLICANT: XU, NANFIE
;  TITLE OF INVENTION: DIFFERENTIALLY EXPRESSED CONIFER cDNAs, AND THEIR USE
;  TITLE OF INVENTION: IN IMPROVING SOMATIC EMBRYOGENESIS
;  FILE REFERENCE: 7648.0023-00
;  CURRENT APPLICATION NUMBER: US/10/651,991
;  CURRENT FILING DATE: 2003-09-02
;  PRIOR APPLICATION NUMBER: 60/239,250
;  PRIOR FILING DATE: 2000-10-11
;  PRIOR APPLICATION NUMBER: 60/260,882
;  PRIOR FILING DATE: 2001-01-12
;  NUMBER OF SEQ ID NOS: 339
;  SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 131
;  LENGTH: 261
;  TYPE: DNA
;  ORGANISM: Pinus taeda
US-10-651-991-131

   Query Match            100.0%;  Score 261;  DB 10;  Length 261;
   Best Local Similarity  100.0%;  Pred. No. 1.3e-63;
   Matches  261; Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     1 AGGTGACCGTACAGCATTATTGATGTTCATTTTGTTGCTGCAAGTTTTTGCAAGTTTTCCGATTCG  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 AGGTGACCGTACAGCATTATTGATGTTCATTTTGTTGCTGCAAGTTTTTGCAAGTTTTCCGATTCG  60

Qy    61 CTGTGAGGCACCGGAAAACGAGATAAGTTGTAAAGTTGCTCGCTGCTGATTGAGGCACGGA  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 CTGTGAGGCACCGGAAAACGAGATAAGTTGTAAAGTTGCTCGCTGCTGATTGAGGCACGGA  120

Qy   121 AAACGAGATAAGTTGTAAAATTTTGCTCGCTGATTTTTGCTGAATATTTCTCACTAT  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AAACGAGATAAGTTGTAAAATTTTGCTCGCTGATTTTTGCTGAATATTTCTCACTAT  180
```

Qy 181 AAAAAGCATTTCCAGAAATAAGAAGGAGCTTTCGAACTGGTTTCCCCAAGAGTTGTAG 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 AAAAAGCATTTCCAGAAATAAGAAGGAGCTTTCGAACTGGTTTCCCCAAGAGTTGTAG 240

Qy 241 GGGGTTTTTCCACGGTCACCT 261
        |||||||||||||||||||||
Db 241 GGGGTTTTTCCACGGTCACCT 261

RESULT 2
US-10-433-793-180
; Sequence 180, Application US/10433793
; Publication No. US20040142334A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von mit Angiogenese assoziierten Krankheiten
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/433,793
; CURRENT FILING DATE: 2003-06-06
; NUMBER OF SEQ ID NOS: 212
; SEQ ID NO 180
; LENGTH: 40324
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-433-793-180

Query Match   14.3%; Score 37.4; DB 8; Length 40324;
Best Local Similarity   52.2%; Pred. No. 31;
Matches   83; Conservative   0; Mismatches   76; Indels   0; Gaps   0;

Qy   25 TGTTCTATTTGTTGTTGCAAGTTTTCCGATTCGCTGTGAGGCACGGAAAACGAGATA 84
          |||| |||| |||||||||||| |||||||||
Db 26359 TGTTTAGTTTTAATATTTTAATTGTTTGTGAAAAGATGAAGATAATTGAGGGAGAGATT 26418

Qy   85 AGTTGTAAAAGTTTGCTCGCTGATTGAGGCACGGAAAACGAGATAAGTTGTAAAATTT 144
          ||||||||| |||
Db 26419 TTTTTAAGAAGTTTGTAGAATGATTTGATTTTAGATAAGTAGTTAGATGTATAAATTAGTTT 26478

Qy   145 GCTCGCTGATTTTTGCTGAATATTTCTCTCACTATAAA 183
          |||||||| |||| ||||||
Db 26479 AATTTAAGATTTTATTGAGATAGGTTGTGTATATATA 26517

RESULT 3
US-09-925-065A-125478
; Sequence 125478, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 125478
; LENGTH: 346
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-125478

Query Match   14.3%; Score 37; DB 4; Length 346;
Best Local Similarity   52.2%; Pred. No. 5;
Matches   82; Conservative   0; Mismatches   75; Indels   0; Gaps   0;

Qy   84 AAGTTGTAAAAGTTTGCTCGCTGATTGAGGCACGGAAAACGAGATAAGTTGTAAAATTT 143
         |||| ||| |||||||||||||||||||||
Db   178 AAACTGAAAGACACTAATCTCTGCCTGAGACCCACAAATCAGATTATCTGTTTTTACCT 237

Qy   144 TGCTCGCTGATTTTTGCTGAATATTTCTCTCACTATAAAAAGCATTTTCCAGAAATAAG 203
         ||| |||
Db   238 GGGTGATTTCTATTCCTCTCATTTTGTGTTCTTTTTATAAAACATCTTTGTTAAATAAT 297

Qy   204 AAGGAGCTTTCGAACTGGTTTCCCCAAGAGTTGTAG 240
         |||| ||
Db   298 AAGTTATTACTCTAAGGTAGTGTCCAAGTCTCTTTG 334

RESULT 4
US-09-925-065A-125480
; Sequence 125480, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 125480
; LENGTH: 346
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-125480

Query Match   14.3%; Score 37; DB 4; Length 346;
Best Local Similarity   52.2%; Pred. No. 5;
Matches   82; Conservative   0; Mismatches   75; Indels   0; Gaps   0;

Qy   84 AAGTTGTAAAAGTTTGCTCGCTGATTGAGGCACGGAAAACGAGATAAGTTGTAAAATTT 143
         |||| ||| |||||||||||||||||||||
Db   178 AAACTGAAAGACACTAATCTCTGCCTGAGACCCACAAATCAGATTATCTGTTTTTACCT 237

Qy   144 TGCTCGCTGATTTTTGCTGAATATTTCTCTCACTATAAAAAGCATTTTCCAGAAATAAG 203
         ||| |||
Db   238 GYGTGATTTCTATTCCTCTCATTTTGTGTTCTTTTTATAAAACATCTTTGTTAAATAAT 297

Qy   204 AAGGAGCTTTCGAACTGGTTTCCCCAAGAGTTGTAG 240
         |||| ||
Db   298 AAGTTATTACTCTAAGGTAGTGTCCAAGTCTCTTTG 334

RESULT 5
US-09-925-065A-125481
; Sequence 125481, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135

QM nucleic - nucleic search, using sw model

```
            GenCore version 5.1.9
    Copyright (c) 1993 - 2006  Bioceleration Ltd.
```

Run on: July 11, 2006, 07:01:45 ; Search time 333.602 Seconds
                                  (without alignments)
                          5454.872 Million cell updates/sec

Title: US-09-973-994-131
Perfect score: 261
Sequence: 1 aggtgaccgcacagcatta.........gggtttttccacggtcacct 261

Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   N_Geneseq_8:*
              1: geneseqn1980s:*
              2: geneseqn1990s:*
              3: geneseqn2000s:*
              4: geneseqn2001s:*
              5: geneseqn2001bs:*
              6: geneseqn2002as:*
              7: geneseqn2002bs:*
              8: geneseqn2003as:*
              9: geneseqn2003bs:*
             10: geneseqn2003cs:*
             11: geneseqn2003ds:*
             12: geneseqn2004as:*
             13: geneseqn2004bs:*
             14: geneseqn2005s:*
             15: geneseqn2006s:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 261 | 100.0 | 261 | 14 | AEA49636 | Aea49636 Loblolly |
| 2 | 37.4 | 14.3 | 40324 | 6 | ABQ67150 | Abq67150 Human ang |
| 3 | 36.8 | 14.1 | 357 | 6 | AAC56859 | Aac56859 Pinus rad |
| 4 | 36.8 | 14.1 | 1943 | 14 | ADW16980 | Adw16980 Pinus rad |
| 5 | 36.8 | 14.1 | 5276 | 4 | AAS46380 | Aas46380 Tumour su |
| 6 | 36.8 | 14.1 | 5276 | 6 | ABL32827 | Abl32827 Human imm |
| 7 | 36.8 | 14.1 | 40862 | 6 | ABL34073 | Abl34073 Human imm |
| 8 | 36.4 | 13.9 | 1084 | 5 | AAC59973 | Aac59973 Human sec |
| 9 | 36 | 13.8 | 37515 | 5 | ABQ66997 | Abq66997 Human ang |
| c 10 | 36 | 13.8 | 2000 | 11 | ACL37108 | Acl37108 Rice stre |
| 11 | 35.4 | 13.6 | 6113 | 11 | ABL32803 | Abl32803 Human imm |
| c 12 | 35.4 | 13.6 | 1726 | 11 | ADO77946 | Ado77946 Brassica |
| 13 | 35 | 13.4 | 73334 | 6 | ABL34124 | Abl34124 Human imm |
| 14 | 35 | 13.4 | 73334 | 6 | ABL92318 | Abl92318 Chemicall |
| 15 | 34.8 | 13.3 | 5542 | 6 | ABL34021 | Abl34021 Human imm |
| 16 | 34.8 | 13.3 | 11209 | 6 | ABN80136 | Abn80136 Human che |
| 17 | 34.6 | 13.3 | 2441 | 14 | ADW10461 | Adw10461 Colon pro |
| 18 | 34.6 | 13.3 | 7135 | 4 | AAS46423 | Aas46423 Tumour su |
| c 19 | 34.4 | 13.2 | 857 | 4 | AAI94399 | Aai94399 Human neu |
| c 20 | 34.4 | 13.2 | 5272 | 13 | ADR07576 | Adr07576 Full leng |
| 21 | 34.4 | 13.2 | 110000 | 2 | AAX20248_06 | Aax20248_06 Continuation (7 of |
| 22 | 34.4 | 13.2 | 111309 | 2 | AAX20250 | Aax20250 Borrelia |
| 23 | 34.2 | 13.1 | 353 | 4 | AAL11902 | Aal11902 Human bre |
| 24 | 34.2 | 13.1 | 362 | 4 | AAL20792 | Aal20792 Human bre |
| 25 | 34.2 | 13.1 | 391 | 4 | AAL11979 | Aal11979 Human bre |
| 26 | 34.2 | 13.1 | 424 | 11 | ACN82095 | Acn82095 Breast ca |
| c 27 | 34.2 | 13.1 | 505 | 11 | ACN90516 | Acn90516 Breast ca |
| 28 | 34.2 | 13.1 | 888 | 11 | ACN82168 | Acn82168 Breast ca |
| 29 | 34.2 | 13.1 | 18683 | 6 | ABL32312 | Abl32312 Human imm |
| 30 | 34.2 | 13.1 | 18683 | 8 | ABL54333 | Abl54333 Chemicall |
| 31 | 34 | 13.0 | 3025 | 8 | ABZ10190 | Abz10190 Haematopo |
| 32 | 34 | 13.0 | 3025 | 10 | ADB54220 | Adb54220 Pretreate |
| 33 | 34 | 13.0 | 3025 | 10 | ADE84158 | Ade84158 Human lym |
| 34 | 34 | 13.0 | 8711 | 6 | AAS46699 | Aas46699 Tumour su |
| 35 | 34 | 13.0 | 13131 | 6 | ABL92249 | Abl92249 Chemicall |
| 36 | 33.8 | 13.0 | 473 | 9 | ACH28682 | Ach28682 Human adu |
| 37 | 33.8 | 13.0 | 562 | 5 | ABV57017 | Abv57017 Human pro |
| 38 | 33.8 | 13.0 | 6545 | 10 | ADE71194 | Ade71194 Novel hum |
| 39 | 33.8 | 13.0 | 37314 | 4 | AAK71358 | Aak71358 Human imm |
| 40 | 33.8 | 13.0 | 172570 | 4 | ABQ88207 | Abq88207 Human ost |
| 41 | 33.6 | 12.9 | 1974 | 4 | AAS60993 | Aas60993 Human can |
| 42 | 33.6 | 12.9 | 2092 | 5 | ABA15073 | Aba15073 Human ner |
| 43 | 33.6 | 12.9 | 3479 | 4 | ABL03600 | Abl03600 Drosophil |
| c 44 | 33.6 | 12.9 | 3968 | 4 | ABL03602 | Abl03602 Drosophil |
| c 45 | 33.6 | 12.9 | 5349 | 6 | ABL32340 | Abl32340 Human imm |

## ALIGNMENTS

```
RESULT 1
AEA49636
ID   AEA49636 standard; cDNA; 261 BP.
XX
AC   AEA49636;
XX
DT   11-AUG-2005  (first entry)
XX
DE   Loblolly pine cDNA clone, LP2-061, SEQ ID 131.
XX
KW   Plant; embryogenesis; plant breeding; paper; wood; ss.
XX
OS   Pinus taeda.
XX
PN   US2005125161-A1.
XX
PD   09-JUN-2005.
XX
PF   02-SEP-2003; 2003US-00651991.
XX
PR   11-OCT-2000; 2000US-0239250P.
PR   12-JAN-2001; 2001US-0260882P.
PR   11-OCT-2001; 2001US-00973994.
XX
PA   (PAPE-) INST PAPER SCI & TECHNOLOGY INC.
XX
PI   Cairney J, Xu N;
XX
DR   WPI; 2005-417092/42.
XX
PT   Relational database of cDNA molecules including those corresponding to
PT   Loblolly pine major intrinsic protein, being differentially expressed
PT   during plant embryogenesis, useful for staging plant embryos.
XX
PS   Claim 17; SEQ ID NO 131; 219pp; English.
XX
CC   The present invention relates to a relational database of cDNA molecules
CC   comprising multiple nucleotide sequences (AEA49506-AEA49832). The
CC   database is useful for staging plant embryos, selecting advantageous
CC   plant clones, determining embryo fitness, and selecting advantageous
CC   conditions for embryo development. The invention is particularly intended
```

CC for use in breeding trees with superior characteristics for use in
CC production of paper and other wood products.
XX
SQ Sequence 261 BP; 69 A; 43 C; 60 G; 89 T; 0 U; 0 Other;

Query Match   100.0%; Score 261; DB 14; Length 261;
Best Local Similarity   100.0%; Pred. No. 1.1e-63;
Matches 261; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 AGGTGACCGTACAGCATTATTGATGTTCTATTTTGTTGCAAGTTTTCCGATTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 AGGTGACCGTACAGCATTATTGATGTTCTATTTTGTTGCAAGTTTTCCGATTCG 60

Qy   61 CTGTGAGGCACGGAAAACGAGATAAGTTGCTCGCTGATTGAGGCACGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTGTGAGGCACGGAAAACGAGATAAGTTGCTCGCTGATTGAGGCACGGA 120

Qy   121 AAACGAGATAAGTTGTAAAATTTTGCTCGCTGAATATTTCTCTCACTAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AAACGAGATAAGTTGTAAAATTTTGCTCGCTGAATATTTCTCTCACTAT 180

Qy   181 AAAAAGCATTTCCAGAAATAAGAAGGAGCTTCGAACTGGTTTCCCCAAGAGTTGTAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AAAAAGCATTTCCAGAAATAAGAAGGAGCTTCGAACTGGTTTCCCCAAGAGTTGTAG 240

Qy   241 GGGGTTTTTCCACGGTCACCT 261
              |||||||||||||||||||||
Db   241 GGGGTTTTTCCACGGTCACCT 261

RESULT 2
ABQ67150
ID   ABQ67150 standard; DNA; 40324 BP.
XX
AC   ABQ67150;
XX
DT   28-AUG-2002 (first entry)
XX
DE   Human angiogenesis associated polynucleotide SEQ ID NO 180.
XX
KW   Human; angiogenesis; methylation; eye disease; glaucoma; tumour;
KW   inflammation; rheumatoid arthritis; diabetic retinopathy; antiulcerds;
KW   macular degeneration; inflammatory bowel disease; Crohn's disease;
KW   antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW   antiarteriosclerotic; ds.
XX
OS   Homo sapiens.
XX
PN   WO20024654-A2.
XX
PD   13-JUN-2002.
XX
PF   06-DEC-2001; 2001WO-EP014320.
XX
PR   06-DEC-2000; 2000DE-01061338.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Schacht O;
XX
DR   WPI; 2002-500450/53.
XX
PT   New nucleic acid fragments from chemically treated angiogenesis-
PT   associated genes, useful for determining methylation status, e.g. in
PT   diagnosis or treatment of cancer.
XX
PS   Claim 1; SEQ ID NO 180; 41pp + Sequence Listing; German.
XX
CC   The invention relates to a nucleic acid (I) comprising a segment of 18
CC   bases of chemically pretreated DNA of angiogenesis-associated genes (II)
CC   having sequences (ABQ66971-ABQ67178) or their complements. (I), also
CC   related oligomers, are used to evaluate the methylation status and/or
CC   single-nucleotide polymorphisms, in angiogenesis-related genes, for

CC diagnosis and treatment of eye diseases, proliferative retinopathy,
CC neovascular glaucoma, solid tumours, inflammation, rheumatoid arthritis,
CC diabetic retinopathy, macular degeneration caused by neovascularisation,
CC psoriasis, arteriosclerosis, inflammatory bowel diseases, ulcers and
CC Crohn's disease. Note: The sequence data for this patent did not form
CC part of the printed specification but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 40324 BP; 12266 A; 436 C; 7652 G; 19970 T; 0 U; 0 Other;

Query Match   14.3%; Score 37.4; DB 6; Length 40324;
Best Local Similarity   52.2%; Pred. No. 4.7;
Matches   83; Conservative   0; Mismatches 76; Indels   0; Gaps   0;

Qy   25 TGTTCTATTTTGTTGTTGCAAGTTTTCCGATTCGCTGTGAGGCACGGAAAACGAGATA 84
              || | |||| | || |||||| | | |||||| ||| || || | ||
Db   26359 TGTTTAGTTTTAATATTTTAATTGTTGTGAAAAGATGAAGATAATTGAGGGAGAGATT 26418

Qy   85 AGTTGTAAAAGTTGTCGCTGATTGAGGCACGGAAAACGAGATAAGTTGTAAAATTTT 144
              | ||| || |||| | | |||| | | | |||| | || ||| || |||
Db   26419 TTTTTAAGAGTTTGTGAGAATGATTGATTTTAGATAGTAAGTTGATGTATAATTAGTTT 26478

Qy   145 GCTCGCTGATTTTTGCTGAATATTTCTCTCACTATAAA 183
              | || ||||||| | ||||||| | | ||| |
Db   26479 AATTTAAGATTTTATTGAGAATAGGTTTGTGTATATATA 26517

RESULT 3
AAC56859
ID   AAC56859 standard; DNA; 357 BP.
XX
AC   AAC56859;
XX
DT   25-JAN-2001 (first entry)
XX
DE   Pinus radiata transcription factor DNA sequence #305.
XX
KW   Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
KW   poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
KW   basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
KW   homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain;
KW   type 2 Cys2His2; CCAAT box element; MYB; ss.
XX
OS   Pinus radiata.
XX
PN   WO200053724-A2.
XX
PD   14-SEP-2000.
XX
PF   09-MAR-2000; 2000WO-US006112.
XX
PR   11-MAR-1999;   99US-00266513.
PR   18-AUG-1999;   99US-0149485P.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
PA   (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI   Wood M,   Mcgrath A,   Shenk MA,   Glenn M;
XX
DR   WPI; 2000-579369/54.
XX
PT   New isolated polynucleotide encoding a plant transcription factor for
PT   producing a plant e.g. a woody plant, preferably eucalyptus or pine,
PT   having modified gene expression or modified activity of a polypeptide.
XX
PS   Claim 1; Page 531; 747pp; English
XX
CC   The present invention relates to novel plant transcription factors from
CC   Eucalyptus grandis or Pinus radiata. The present sequence is the coding
CC   sequence for one such transcription factor. The transcription factor may
CC   be used to produce a plant having modified gene expression such as a
CC   woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or
CC   mahogany species or to modify the activity of a polypeptide in a plant.

```
                    GenCore version 5.1.9
          Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on:     July 11, 2006, 07:10:56 ; Search time 2616.43 Seconds
                                     (without alignments)
                                     5578.184 Million cell updates/sec

Title:          US-09-973-994-131
Perfect score:  261
Sequence:       1 aggtgaccgtacagcatta.........gggtttttccacggtcacct 261

Scoring table:  IDENTITY.NUC
                Gapop 10.0 , Gapext 1.0

Searched:       48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters:    96473596

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      EST:*
                1:  gb_est1:*
                2:  gb_est3:*
                3:  gb_est4:*
                4:  gb_est5:*
                5:  gb_est6:*
                6:  gb_htc:*
                7:  gb_est2:*
                8:  gb_est7:*
                9:  gb_est8:*
               10:  gb_est9:*
               11:  gb_gss1:*
               12:  gb_gss2:*
               13:  gb_gss3:*
               14:  gb_gss4:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	90.2	34.6	563	13	CZ895441	CZ895441 upta001f0
c	2	89.8	34.4	674	9	DR053876	DR053876 RTCA1_13_
	3	82.4	31.6	817	9	DR120385	DR120385 RTMG1_29_
	4	76	29.1	643	13	CZ895986	CZ895986 226_2_123
c	5	42.2	16.2	1257	7	BB894011	BB894011 601437812
c	6	41.8	16.0	495	8	CO456808	CO456808 MZCCS1500
c	7	40.2	15.4	1101	14	CNS000D1	AL065414 Drosophil
c	8	40.2	15.4	1356	10	DV779559	DV779559 Hw_Fat_57
c	9	40	15.3	783	14	BX197330	BX197330 Danio rer
c	10	39.2	15.0	607	12	CE516027	CE516027 tigr-gss-
	11	38.8	14.9	694	14	BX158594	BX158594 Danio rer
c	12	38.8	14.9	941	3	BU153366	BU153366 AGENCOURT
c	13	38.8	14.8	386	8	CN252666	CN252666 EST018601
c	14	38.6	14.8	663	14	DE073250	DE073250 Oryzias_l
c	15	38.6	14.8	1101	14	CNS00EQL	AL069526 Drosophil
c	16	38.6	14.8	1101	14	CNS00ZB7	AL097453 Drosophil
c	17	38.4	14.7	489	4	BX553089	BX553089 BX553089
c	18	38.4	14.7	856	8	CV595856	CV595856 L_AJ-aaa0
c	19	38.2	14.6	537	12	CE105632	CE105632 tigr-gss-

c	20	38.2	14.6	677	11	BZ278020	BZ278020 CH230-356
	21	38.2	14.6	792	11	B2455713	B2455713 BONSH83TF
	22	37.8	14.5	603	10	DR783465	DR783465 Le_ivo_17
	23	37.8	14.5	610	10	DR714239	DR714239 Le_ivo_13
	24	37.8	14.5	620	8	CO049664	CO049664 Le_mx0_03
c	25	37.8	14.5	830	11	AQ750770	AQ750770 HS_5576_B
	26	37.8	14.4	1505	5	AM102256	AM102256
	27	37.6	14.4	691	3	BP677423	BP677423
c	28	37.4	14.4	1322	10	DV795774	DV795774 Hw_loin_3
c	29	37.4	14.3	620	9	CX506668	CX506668 JGI_XG51
c	30	37.4	14.3	866	14	DU738879	DU738879 APK14906.
c	31	37.4	14.3	908	14	CNS01678	AJ106382 Drosophil
c	32	37.2	14.3	299	14	DE245139	DE245139 Trifolium
c	33	37.2	14.3	531	7	AW991228	AW991228 RDEV151_H
	34	37.2	14.3	650	12	CC847241	CC847241 NDL.109D1
	35	37.2	14.3	650	12	CC853143	CC853143 NDL.109C1
	36	37	14.2	448	2	BI275009	BI275009 UI-R-CX0-
	37	37	14.2	663	2	BI284287	BI284287 UI-R-CX0s
c	38	37	14.2	682	14	BX195327	BX195327 Danio rer
	39	37	14.2	940	13	DU045365	DU045365 153078 To
c	40	36.8	14.1	788	14	DU591489	DU591489 OO_Ba008
c	41	36.8	14.1	845	14	DU865836	DU865836 82515 Tom
c	42	36.8	14.1	910	14	DU875609	DU875609 147915 To
c	43	36.6	14.0	963	13	DU082579	DU082579 271367 To
	44	36.6	14.0	618	12	CE048040	CE048040 tigr-gss-
	45	36.6	14.0	672	11	BH514532	BH514532 BOHQL25TF

```
                                 ALIGNMENTS

RESULT 1
CZ895441/c
LOCUS       CZ895441              563 bp    DNA      linear     GSS 15-SEP-2005
DEFINITION  upta001f002b10f0 Pine methylation unfiltered library (LibID: 131)
            Pinus taeda genomic, genomic survey sequence.
ACCESSION   CZ895441
VERSION     CZ895441.1  GI:75689927
KEYWORDS    GSS.
SOURCE      Pinus taeda (loblolly pine)
  ORGANISM  Pinus taeda
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
REFERENCE   1  (bases 1 to 563)
  AUTHORS   Rabinowicz,P.D., Citek,R.W., Budiman,M.A., Nunberg,A., Bedell,J.A.,
            Lakey,N., O'Shaughnessy,A.L., Nascimento,L.U., McCombie,W.R. and
            Martienssen,R.A.
  TITLE     Differential methylation of genes and repeats in land plants
  JOURNAL   Genome Res. 15 (10), 1431-1440 (2005)
  PUBMED    16204196
COMMENT     Contact: Rabinowicz PD
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 795 7787
            Fax: 301 838 0208
            Email: pablo@tigr.org
            Class: methylation filtered
            High quality sequence stop: 563.
FEATURES             Location/Qualifiers
     source          1..563
                     /organism="Pinus taeda"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3352"
                     /clone_lib="Pine methylation unfiltered library (LibID:
                     131)"
ORIGIN

Query Match        34.6%;   Score 90.2; DB 13; Length 563;
Best Local Similarity   88.9%;   Pred. No. 3e-13;
Matches 120;  Conservative 0; Mismatches 13; Indels 2; Gaps 2;

Qy    127 GATAAGTGTAAATTTGCTCGCTGATTTTTGCTGAATATTTCTCCACTATAAAAG 186
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
              unidirectionally into pSL1180. Inserts can be excised with
              EcoRI (5' end) and XhoI (3' end)."

ORIGIN

Query Match        34.4%;  Score 89.8; DB 9; Length 674;
Best Local Similarity 83.2%;  Pred. No. 3.9e-13;
Matches 114;  Conservative  0;  Mismatches 22; Indels 1; Gaps 1;

Qy  125 GAGATAAGTTGTAAATTTTGCTCGCTGATTTTTTGCTGAATATTTCTCACTATAAA 184
         ||||||||||||||| |||||| ||||||  |||||||||||||||||| |||||||
Db  511 GAGATAAGTTGTAAAAGTTGCTTGCTGA-TTTTTGCTGAATATTTCTTGCTGTGTAAAA 453

Qy  185 AGCATTTTCCAGAAATAAGAGGAGCGTTTCGAACTGGTTTTCCCCAAGAGTTGTAGGGGG 244
         ||||||| |||||||||| ||| |||||| || |||||||||||||||||||||||||||
Db  452 GACATTTTTCAGAAATAAAGAGAAGAGCTCTCAAACTGGTTTTCCCCAAGAGTTGTAGGGG 393

Qy  245 TTTTCCACGGTCACCT 261
         ||||| | |||||||
Db  392 TATTTCCAGGGTCACAT 376


RESULT 3
DR120385
LOCUS       DR120385                817 bp    mRNA            linear   EST 14-JUN-2005
DEFINITION  RTMG1_29_H05.b1_A029 Roots minus magnesium Pinus taeda cDNA clone
            RTMG1_29_H05_A029 3', mRNA sequence.
ACCESSION   DR120385
VERSION     DR120385.1  GI:67708695
KEYWORDS    EST.
SOURCE      Pinus taeda (loblolly pine)
  ORGANISM  Pinus taeda
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
REFERENCE   1 (bases 1 to 817)
  AUTHORS   Pratt,L., Cordonnier-Pratt,M.M., Lorenz,W.W., Zimmermann,C. and
            Dean,J.F.D.
  TITLE     An EST database from magnesium-deficient loblolly pine (Pinus
            taeda) roots
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Cordonnier-Pratt MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            RNA prepared and library constructed by W. Walter Lorenz (School of
            Forest Resources, University of Georgia); plant material prepared
            by Craig Zimmermann (School of Forest Resources, University of
            Georgia) using rooted cuttings provided by the Forest Biology
            Research Cooperative (FBRC) and the CCLONES project a the
            University of Florida; sequencing done in the Laboratory for
            Genomics and Bioinformatics, University of Georgia. Sequence ends
            have been trimmed to exclude vector and regions below Phred quality
            16. Three-prime sequences are presented as their reverse complement
            and have been trimmed to exclude polyA.
            Seq primer: M13-21 (TGTAAAACGACGGCCAGT)
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1..817
                     /organism="Pinus taeda"
                     /mol_type="mRNA"
                     /strain="3 CCLONES"
                     /db_xref="taxon:3352"
                     /clone="RTMG1_29_H05_A029"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="Roots minus magnesium"
                     /note="Organ: Root; Vector: pSL1180; Site_1: EcoRI;
                     Site_2: XhoI; The library was prepared from polyA+ RNA
                     from the roots of 1-year-old loblolly pine (Pinus taeda)
                     cuttings that were rooted and then planted in washed sand.
                     The rooted cuttings were maintained for 139 days (July 28
                     2003 harvest) under ambient conditions in a local
```

```
Db  563 GATAAGTTGTAAAAGTTGCTTGCTGCTGATTTTTTGCTGA-ATATTCTCACTATAAAAAG 505

Qy  187 CATTTTTCCAGAAATAAGAGGAGCGTTTCGAACTGGTTTTCCCCAAGAGTTGTAGGGGTT 246
         ||||||  |||||||||| || |||||  |||||||||||| |||||||||||||||||
Db  504 CATTTT-CAGAAATAAGAGAGGAGCGAGCTCCGAACTGGTTTTCTCCAAGAGTTGTAGGGGGTT 446

Qy  247 TTTCCACGGTCACCT 261
         |||| | ||||||||
Db  445 TTTTCAGGGTCACCT 431


RESULT 2
DR053876/c
LOCUS       DR053876                674 bp    mRNA            linear   EST 06-JUN-2005
DEFINITION  RTCA1_13_H01.g1_A029 Roots minus calcium Pinus taeda cDNA clone
            RTCA1_13_H01_A029 5', mRNA sequence.
ACCESSION   DR053876
VERSION     DR053876.1  GI:66977443
KEYWORDS    EST.
SOURCE      Pinus taeda (loblolly pine)
  ORGANISM  Pinus taeda
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
REFERENCE   1 (bases 1 to 674)
  AUTHORS   Pratt,L., Cordonnier-Pratt,M.M., Lorenz,W.W., Zimmermann,C. and
            Dean,J.F.D.
  TITLE     An EST database from calcium-deficient loblolly pine (Pinus taeda)
            roots
  JOURNAL   Unpublished (2005)
COMMENT     Other ESTs: RTCA1_13_H01.b1_A029
            Contact: Cordonnier-Pratt MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            RNA prepared and library constructed by W. Walter Lorenz (School of
            Forest Resources, University of Georgia); plant material prepared
            by Craig Zimmermann (School of Forest Resources, University of
            Georgia) using rooted cuttings provided by the Forest Biology
            Research Cooperative (FBRC) and the CCLONES project a the
            University of Florida; sequencing done in the Laboratory for
            Genomics and Bioinformatics, University of Georgia. Sequence ends
            have been trimmed to exclude vector and regions below Phred quality
            16. Three-prime sequences are presented as their reverse complement
            and have been trimmed to exclude polyA.
            Seq primer: JENREV (CAGGAAACAGCTATGACC).
FEATURES             Location/Qualifiers
     source          1..674
                     /organism="Pinus taeda"
                     /mol_type="mRNA"
                     /strain="3 CCLONES"
                     /db_xref="taxon:3352"
                     /clone="RTCA1_13_H01_A029"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="Roots minus calcium"
                     /note="Organ: Root; Vector: pSL1180; Site_1: EcoRI;
                     Site_2: XhoI; The library was prepared from polyA+ RNA
                     from the roots of 1-year-old loblolly pine (Pinus taeda)
                     cuttings that were rooted and then planted in washed sand.
                     The rooted cuttings were maintained for 139 days (July 28
                     2003 harvest) under ambient conditions in a local
                     greenhouse. They were kept on a weekly regimen of 0.5x
                     nutrient-complete Hoagland's solution and supplemented
                     with additional water sufficient to maintain a 15% soil
                     moisture content. Starting five weeks prior to harvesting
                     roots for mRNA preparation, the weekly fertilizer regimen
                     for the potted trees was modified to omit calcium from the
                     0.5x Hoagland's solution (5 applications). For two days
                     prior to starting the modified fertilizer regimen, pots
                     were flushed extensively with water to reduce residual
                     levels of calcium. Double-stranded cDNA was cloned
```

```
               Copyright   GenCore version 5.1.9
               Copyright (c) 1993 - 2006  Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on:   July 11, 2006, 07:02:26 ; Search time 1779.21 Seconds
                                        (without alignments)
                            9380.717 Million cell updates/sec

Title:          US-09-973-994-131
Perfect score:  261
Sequence:       1 aggtgaccgtacagcatta.........gggtttttccacggtcacct 261

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       6366136 seqs, 3197371052 5 residues

Total number of hits satisfying chosen parameters:   12732272

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
              1: gb_env:*
              2: gb_pat:*
              3: gb_ph:*
              4: gb_pl:*
              5: gb_pr:*
              6: gb_ro:*
              7: gb_sts:*
              8: gb_sy:*
              9: gb_un:*
             10: gb_vi:*
             11: gb_ov:*
             12: gb_htg:*
             13: gb_in:*
             14: gb_om:*
             15: gb_ba:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	Query Match	Length	DB	ID	Description
c	1	42	16.1	171009	12	AC128851	AC128851 Rattus no
c	2	42	16.0	264832	12	AC106309	AC106309 Rattus no
c	3	41.8	16.0	38744	13	CEC56A3	Z77655 Caenorhabdi
	4	41.2	15.8	161505	6	AC154234	AC154234 Mus muscu
	5	41.2	15.6	222907	12	CT009561	CT009561 Mus muscu
	6	40.8	15.6	182957	12	CR931799	CR931799 Danio rer
	7	40.8	15.6	259146	12	CT573231	CT573231 Danio rer
c	8	40.6	15.6	231271	12	AC117004	AC117004 Rattus no
c	9	40.6	15.6	232508	12	AC116189	AC116189 Rattus no
c	10	39.8	15.2	237705	12	AC171299	AC171299 Bos tauru
c	11	39.8	15.2	239162	12	AC175441	AC175441 Bos tauru
c	12	39.4	15.1	94646	12	AC171129	AC171129 Helobdell
c	13	39.2	15.0	158135	12	AC080040	AC080040 Homo sapi
	14	39.2	15.0	174058	12	AC068695	AC068695 Homo sapi
c	15	39.2	15.0	176570	5	AC009499	AC009499 Homo sapi
	16	38.8	14.9	33651	13	AC114261	AC114261 Dictyoste
	17	38.8	14.9	195124	11	CR548625	CR548625 Zebrafish
	18	38.8	14.9	208436	12	CT027756	CT027756 Danio rer
	19	38.6	14.8	7942	10	HPV52	X74481 Human papil
	20	38.6	14.8	155399	5	PT027015	CR937025 Pan trogl
c	21	38.6	14.8	192290	12	AC153218	AC153218 Bos tauru
c	22	38.6	14.8	195220	12	AC151020	AC151020 Callithri
	23	38.6	14.8	249791	11	BX537133	BX537133 Zebrafish
	24	38.6	14.6	271777	12	AC157148	AC157148 Bos tauru
	25	38.2	14.6	103409	11	AL662880	AL662880 Zebrafish
	26	38.2	14.6	163804	11	BX005463	BX005463 Zebrafish
	27	38.2	14.6	170688	6	AC101795	AC101795 Mus muscu
c	28	38.2	14.6	222137	12	CR954958	CR954958 Danio rer
c	29	38.2	14.6	281676	12	AC171784	AC171784 Bos tauru
c	30	38	14.6	35850	13	AC006791	AC006791 Caenorhab
	31	38	14.6	230810	12	AC118784	AC118784 Rattus no
c	32	37.8	14.5	180309	5	AC073576	AC073576 Homo sapi
	33	37.8	14.5	188107	12	CNS01DWE	AL137129 Human chr
	34	37.8	14.5	264773	12	AC163148	AC163148 Bos tauru
	35	37.6	14.4	73250	5	AC114946	AC114946 Homo sapi
	36	37.6	14.4	114214	12	AP007470	AP007470 Lotus cor
	37	37.6	14.4	168419	12	AC025384	AC025384 Homo sapi
c	38	37.6	14.4	224828	12	AC157113	AC157113 Bos tauru
	39	37.4	14.3	40324	2	AX458634	AX458634 Sequence
c	40	37.4	14.3	110000	12	PFMAL8P1_04	Continuation (5 of
c	41	37.4	14.3	125632	4	AC015446	AC015446 Arabidops
	42	37.4	14.3	146130	12	AC015875	AC015875 Homo sapi
	43	37.4	14.3	200195	5	AC146220	AC146220 Pan trogl
c	44	37.2	14.3	27426	5	AC079740	AC079740 Homo sapi
c	45	37.2	14.3	132822	12	AC181755	AC181755 Strongylo

ALIGNMENTS

```
RESULT 1
AC128851/c
LOCUS       AC128851         171009 bp    DNA     linear   HTG 20-NOV-2002
DEFINITION  Rattus norvegicus clone CH230-445E24, WORKING DRAFT SEQUENCE.
ACCESSION   AC128851
VERSION     AC128851.3  GI:25139471
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 171009)
AUTHORS     Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,C., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,J.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,R., Garcia,M., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,T., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozano,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Meenen,E.,
```

```
                                                      * by the finished sequence as soon as it is available and
                                                      * the accession number will be preserved.
                                                      *
                                          FEATURES             Location/Qualifiers
                                          source           1..171009
                                                            /organism="Rattus norvegicus"
                                                            /mol_type="genomic DNA"
                                                            /db_xref="taxon:10116"
                                                            /clone="CH230-445E24"
                                          misc_feature     1..2653
                                                            /note="wgs end extension
                                                            clone end:Sp6"
                                          misc_feature     5309..6432
                                                            /note="clone boundary
                                                            clone end:Sp6
                                                            site:
                                                            end_sequence:BZ201009"
                                          misc_feature     167893..168564
                                                            /note="clone boundary
                                                            clone end:T7
                                                            site:
                                                            end_sequence:BZ201007"
                                          ORIGIN

Query Match        16.1%;   Score 42;   DB 12;   Length 171009;
Best Local Similarity  47.7%;   Pred. No. 0.019;
Matches 123;  Conservative  0;  Mismatches 135;  Indels  0;  Gaps  0;

Qy      4 TGACCGTACAGCATTATTGATGTCTATTTGTGTTGTTGCAAGTTTTCCGATTCGCTG 63
          ||||||||||||||||||||||||||||||||||| |||||| || ||||| |||||
Db 123083 TGGCGGAGCAACCAGAGCTCATGGTCGTGTAGTAGTTGTGATATGTTCACCGTTGGTGGGAGTTCATG 123024

Qy     64 TGAGGCACGGAAAACAGAGATAAGTTGTAAAAGTTGGCTCGCTGATTGAGGCACGGAAAA 123
          |||||||||| ||||  |||||| ||||||||  |||| |||||||||||||||| ||||
Db 123023 TGCTCCTAACAAAATGATAAAACTGAAAGAATATGTTAGCTTTATTTAATTAACCATTGCAAA 122964

Qy    124 CGAGATAAGTTGTAAAATTTTGCTCGCTGATTTTTGCTGAATATTTCTCTACTATAAA 183
          ||||| |||||||  |||||||| |||||| |||||   |||   |||||||||||||
Db 122963 TTTTACAAATTGTATGTGTAAGATGCACTTTTTACCTCTAAAATATATTTGGACATAATTT 122904

Qy    184 AAGCATTTTCCAGAAATAAAGAAGAGGCTTTCGAACTGGTTTCCCCAAGAGTTGTAGGGG 243
          |||||||| |||   ||| ||| |  ||||||||||  |||| |||||||||||  ||||
Db 122903 GTCAATTTATATATTTTAAGAAAAATACAGATCTTAGACTATTAGTTGTCAAGA 122844

Qy    244 GTTTTTCCACGGTCACCT 261
          |||| |||||||| ||||
Db 122843 TTATTAAAAGCAGCTCCT 122826


RESULT 2
AC106309
LOCUS       AC106309              264832 bp    DNA     linear   HTG 13-MAY-2003
DEFINITION  Rattus norvegicus clone CH230-121P20, WORKING DRAFT SEQUENCE, 4
            unordered pieces.
ACCESSION   AC106309
VERSION     AC106309.4  GI:30581468
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 264832)
AUTHORS     Muzny,D.Marie, Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Albbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
```

```
                          Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
                          Morgan,M., Morris,G., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
                          Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
                          Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
                          Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
                          Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
                          Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
                          Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
                          Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
                          Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
                          Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
                          Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
                          Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
                          Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
                          Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
                          Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
                          Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
                          Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
                          Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
                          Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
                          Weinstock,G. and Gibbs,R.A.
TITLE       Direct Submission
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 171009)
AUTHORS     Worley,K.C.
TITLE       Direct Submission
JOURNAL     Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 171009)
AUTHORS     Rat Genome Sequencing Consortium.
TITLE       Direct Submission
JOURNAL     Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Nov 20, 2002 this sequence version replaced gi:23914598.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            --------- Genome Center
            Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu/
            Contact: hgsc-help@bcm.tmc.edu
            --------- Project Information
            Center project name: KAVY
            Center clone name: CH230-445E24
            --------- Summary Statistics
            Assembly program: Phrap; version 0.990329
            Consensus quality: 164467 bases at least Q40
            Consensus quality: 166274 bases at least Q30
            Consensus quality: 167295 bases at least Q20
            Estimated insert size: 171470; sum-of-contigs estimation
            Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            *   consists of 1 contigs. Gaps between the contigs
            *   are represented as runs of N. The order of the pieces
            *   is believed to be correct as given, however the sizes
            *   of the gaps between them are based on estimates that have
            *   provided by the submittor.
            * This sequence will be replaced
```

```
                    GenCore version 5.1.9
          Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on:          July 11, 2006, 07:21:26 ; Search time 125.935 Seconds
                                           (without alignments)
                                           4561.330 Million cell updates/sec

Title:           US-09-973-994-79
Perfect score:   307
Sequence:        1 ggtggcgatcctagaattgca..........gccatcactccggatcgcacc 307

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters:        2807332

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued Patents NA:*
                 1: /EMC_Celerra_SIDS3/ptodata/2/ina/1_COMB.seq:*
                 2: /EMC_Celerra_SIDS3/ptodata/2/ina/5_COMB.seq:*
                 3: /EMC_Celerra_SIDS3/ptodata/2/ina/6A_COMB.seq:*
                 4: /EMC_Celerra_SIDS3/ptodata/2/ina/6B_COMB.seq:*
                 5: /EMC_Celerra_SIDS3/ptodata/2/ina/7_COMB.seq:*
                 6: /EMC_Celerra_SIDS3/ptodata/2/ina/H_COMB.seq:*
                 7: /EMC_Celerra_SIDS3/ptodata/2/ina/PCTUS_COMB.seq:*
                 8: /EMC_Celerra_SIDS3/ptodata/2/ina/PP_COMB.seq:*
                 9: /EMC_Celerra_SIDS3/ptodata/2/ina/RE_COMB.seq:*
                 10: /EMC_Celerra_SIDS3/ptodata/2/ina/backfiles1.seq:*

       Pred. No. is the number of results predicted by chance to have a
       score greater than or equal to the score of the result being printed,
       and is derived by analysis of the total score distribution.

                                 SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description
1	55.8	18.2	7218	2	US-08-232-463-14	Sequence 14, Appl
2	36.4	11.9	63588	3	US-09-873-404-3	Sequence 3, Appli
3	36.4	11.9	63588	3	US-10-243-735-3	Sequence 3, Appli
4	36.4	11.9	63588	4	US-10-730-010-3	Sequence 3, Appli
5	36	11.7	219964	3	US-09-949-016-15086	Sequence 15086, A
6	35.4	11.5	101349	3	US-09-949-016-17433	Sequence 17433, A
c 7	34.8	11.3	1664976	3	US-08-916-421B-1	Sequence 1, Appli
c 8	34.8	11.3	1664976	3	US-09-692-570-1	Sequence 1, Appli
c 9	34.6	11.3	5394	3	US-08-688-376-1	Sequence 1, Appli
10	34.4	11.2	764	3	US-09-949-016-5644	Sequence 5644, Ap
11	34.4	11.2	764	3	US-09-949-016-5645	Sequence 5645, Ap
c 12	34.4	11.2	1127	3	US-09-976-594-1102	Sequence 1102, Ap
13	34	11.1	15756	3	US-09-949-016-13607	Sequence 13607, A
14	34	11.1	29165	3	US-09-949-016-12340	Sequence 12340, A
15	33.8	11.0	1141	2	US-09-806-708B-22	Sequence 22, Appl
c 16	33.4	10.9	453	2	US-09-328-352-2667	Sequence 2667, Ap
c 17	33.4	10.9	601	3	US-09-949-016-64949	Sequence 64949, A
18	33.4	10.9	87863	3	US-09-949-016-14402	Sequence 14402, A
19	33.4	10.9	300402	3	US-09-949-016-13632	Sequence 13632, A
c 20	33	10.7	396	3	US-09-495-050A-98	Sequence 98, Appl
c 21	33	10.7	6854	2	US-08-468-036-4	Sequence 4, Appli
c 22	33	10.7	6854	2	US-08-376-843-4	Sequence 4, Appli
c 23	33	10.7	24740	3	US-09-949-016-13528	Sequence 13528, A

						ID	Description
c	24	33	32.6	10.7	74962	3	US-09-685-853A-3	Sequence 3, Appli
	25	32.6	10.6	1377	3	US-09-583-110-2506	Sequence 2506, Ap
	26	32.6	10.6	1392	3	US-09-107-433-903	Sequence 903, App
c	28	32.2	10.5	39318	3	US-09-949-016-13798	Sequence 13798, A
c	29	32.2	10.5	601	3	US-09-949-016-64948	Sequence 64948, A
c	30	32.2	10.5	166698	3	US-09-949-016-152910	Sequence 152910,
	31	32	10.4	601	3	US-09-949-016-16038	Sequence 16038, A
	32	32	10.4	601	3	US-09-949-002-3310	Sequence 3310, Ap
	33	32	10.4	601	3	US-09-949-002-3311	Sequence 3311, Ap
	34	32	10.4	601	3	US-09-949-002-7769	Sequence 7769, Ap
	35	32	10.4	51905	3	US-09-949-002-7770	Sequence 7770, Ap
	36	32	10.4	51905	3	US-09-949-002-667	Sequence 667, App
c	37	31.8	10.4	228	3	US-09-949-002-781	Sequence 781, App
c	38	31.8	10.4	601	3	US-09-949-248-796A-11240	Sequence 11240, A
c	39	31.8	10.4	601	3	US-09-949-016-152908	Sequence 152908,
c	40	31.8	10.4	1827	3	US-09-949-016-152909	Sequence 152909,
	41	31.8	10.4	1827	3	US-09-270-767-1308	Sequence 1308, Ap
c	42	31.8	10.4	250352	3	US-09-270-767-16590	Sequence 16590, A
	43	31.6	10.3	100848	3	US-09-596-002-39	Sequence 39, Appl
c	44	31.6	10.3	312474	3	US-09-949-016-17434	Sequence 17434, A
c	45	31.4	10.2	601	3	US-09-949-016-128705	Sequence 128705,

```
                                                               ALIGNMENTS

RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
;   APPLICANT: DORNER, F.
;   APPLICANT: SCHEIFLINGER, F.
;   APPLICANT: FALKNER, F. G.
;   TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
;   NUMBER OF SEQUENCES: 52
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Foley & Lardner
;     STREET: 1800 Diagonal Road, Suite 500
;     CITY: Alexandria
;     STATE: VA
;     COUNTRY: USA
;     ZIP: 22313-0299
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/232,463
;     FILING DATE:
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US/07/935,313
;     FILING DATE:
;     APPLICATION NUMBER: EP 91 114 300.6
;     FILING DATE: 26-AUG-1991
;   ATTORNEY/AGENT INFORMATION:
;     NAME: BENT, Stephen A.
;     REGISTRATION NUMBER: 29,768
;     REFERENCE/DOCKET NUMBER: 3047.2/114 IMMU
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (703)836-9300
;     TELEFAX: (703)683-4109
;     TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 7218 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   IMMEDIATE SOURCE:
```

```
;   CLONE: pTZgpt-F1s
US-08-232-463-14

Query Match         18.2%;   Score 55.8; DB 2; Length 7218;
Best Local Similarity   2.0%;   Pred. No. 1.2e-06;
Matches    6;   Conservative 186;   Mismatches 103;   Indels    0;   Gaps    0;

Qy      3 TGCGATCCTAGAATTGCATCTGTTGCCGTGCTACTCCCTTGAATTAGAAGCCTTCCTT  62
          |||||::::       ::  :::: :::::YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY
Db   1061 TGCGATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1120

Qy     63 GCCACCTGGTCTTGAAGCATCCTTTTAGTCTTATTATCCTTGGCATTATTCCTAGAATCC  122
          :: ::::::YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY
Db   1121 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1180

Qy    123 TCTTGCTCCTTAACATCCTATAAGGTCATCTCATGCATGCCAAGAGATCTTGACAAATTTGGAA  182
          :: ::::::YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY
Db   1181 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1240

Qy    183 GAATCTGCCAATTGCTCCCGCCCTCCTTAACTCGAAGTGACTCTCTAAGTGCTGAACC  242
          :: :::::: ::::::YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY
Db   1241 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1300

Qy    243 TTCTGAGCTACCTGTTCTTTGTTTCTATCATTGCTTTTCTTTGTGCCATCACTC  297
          :: :: :::YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY
Db   1301 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1355

RESULT 2
US-09-873-404-3
; Sequence 3, Application US/09873404
; Patent No. 6500656
GENERAL INFORMATION:
APPLICANT: WEBSTER, Marion et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: CL001212-CIP
CURRENT APPLICATION NUMBER: US/09/873,404
CURRENT FILING DATE: 2001-06-05
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 3
LENGTH: 63588
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(63588)
OTHER INFORMATION: n = A,T,C or G
US-09-873-404-3

Query Match         11.9%;   Score 36.4; DB 3; Length 63588;
Best Local Similarity  51.2%;   Pred. No. 1.6;
Matches   85;   Conservative    0;   Mismatches   81;   Indels    0;   Gaps    0;

Qy    106 CATTATTCCTAGAATCCTCTTGCTCCTTAACATCCTATAAGGTCATCTCATGCCAAGAGA  165
          ||| ||       ||||       |||||
Db  15012 CAGTCTTACTTGCATGCGCTTTCTCCTCACCCTCCAGCCTACAGACTTCTTCACAGCACT 15071

Qy    166 TCTTGACAAATTGGAAGAATCTGCCAATTGCTCCCGCCCTCCTTAACTCGAAGTTGACT  225
          |||| ||       ||||       |||||
Db  15072 TCTCACCCAATGCTGAGGGACTCCCCAACTAGACTTCGCCTCCTTAACAGTAGAATGGAT 15131

Qy    226 CTCTAAGTTGCTGAACCTTCTGAACCTTCTTTTATTCCTACCTGTTCTTTGTTTTTCTAT  271
          |||| ||       ||||       |||||
Db  15132 GGCTTCTTGCTTTGCTTTTATTCCTACACAGCAGCATTCCTTGCTTTGCAT 15177
```

```
; GENERAL INFORMATION:
;   APPLICANT: WEBSTER, Marion et al
;   TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
;   TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
;   TITLE OF INVENTION: THEREOF
;   FILE REFERENCE: CL001212DIV
;   CURRENT APPLICATION NUMBER: US/10/243,735
;   CURRENT FILING DATE: 2002-09-16
;   NUMBER OF SEQ ID NOS: 4
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 63588
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(63588)
;   OTHER INFORMATION: n = A,T,C or G.
US-10-243-735-3

Query Match         11.9%;   Score 36.4; DB 3; Length 63588;
Best Local Similarity  51.2%;   Pred. No. 1.6;
Matches   85;   Conservative    0;   Mismatches   81;   Indels    0;   Gaps    0;

Qy    106 CATTATTCCTAGAATCCTCTTGCTCCTTAACATCCTATAAGGTCATCTCATGCCAAGAGA  165
          ||| ||       ||||       |||||
Db  15012 CAGTCTTACTTGCATGCGCTTTCTCCTCACCCTCCAGCCTACAGACTTCTTCACAGCACT 15071

Qy    166 TCTTGACAAATTGGAAGAATCTGCCAATTGCTCCCGCCCTCCTTAACTCGAAGTTGACT  225
          |||| ||       ||||       |||||
Db  15072 TCTCACCCAATGCTGAGGGACTCCCCAACTAGACTTCGCCTCCTTAACAGTAGAATGGAT 15131

Qy    226 CTCTAAGTTGCTGAACCTTCTGAACCTTCTTTTATTCCTACCTGTTCTTTGTTTTTCTAT  271
          |||| ||       ||||       |||||
Db  15132 GGCTTCTTGCTTTGCTTTTATTCCTACACAGCAGCATTCCTTGCTTTGCAT 15177

RESULT 4
US-10-730-010-3
; Sequence 3, Application US/10730010
; Patent No. 6984511
GENERAL INFORMATION:
APPLICANT: WEBSTER, Marion et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: CL001212CIPDIV-II
CURRENT APPLICATION NUMBER: US/10/730,010
CURRENT FILING DATE: 2003-12-09
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 3
LENGTH: 63588
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(63588)
OTHER INFORMATION: n = A,T,C or G.
US-10-730-010-3

Query Match         11.9%;   Score 36.4; DB 4; Length 63588;
Best Local Similarity  51.2%;   Pred. No. 1.6;
Matches   85;   Conservative    0;   Mismatches   81;   Indels    0;   Gaps    0;

Qy    106 CATTATTCCTAGAATCCTCTTGCTCCTTAACATCCTATAAGGTCATCTCATGCCAAGAGA  165
          ||| ||       ||||       |||||
Db  15012 CAGTCTTACTTGCATGCGCTTTCTCCTCACCCTCCAGCCTACAGACTTCTTCACAGCACT 15071

Qy    166 TCTTGACAAATTGGAAGAATCTGCCAATTGCTCCCGCCCTCCTTAACTCGAAGTTGACT  225
          |||| ||       ||||       |||||
Db  15072 TCTCACCCAATGCTGAGGGACTCCCCAACTAGACTTCGCCTCCTTAACAGTAGAATGGAT 15131
```

OM nucleic - nucleic search, using sw model

           Copyright   GenCore version 5.1.9
                        Copyright (c) 1993 - 2006 Bioacceleration Ltd.

Run on: July 11, 2006, 07:46:32 ; Search time 916.676 Seconds
                                     (without alignments)
                                     4115.199 Million cell updates/sec

Title: US-09-973-994-79
Perfect score: 307
Sequence: 1 ggtgcgatccagaattgca.........gccatcactggatcgcacc 307

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Published_Applications_NA_Main:*
     1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq:*
     2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq:*
     3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
     4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
     5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
     6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
     7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
     8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
     9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
    10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
    11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
    12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
    13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
    14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
    15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq:*
    16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq:*
        /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                                    SUMMARIES

                      %
Result   Query
No.   Score Match Length DB ID   Description
---   ----- ----- ------ -- --   -----------
  1   307   100.0 307   10 US-10-651-991-79   Sequence 79, Appl
  2   296   96.4   308   10 US-10-651-991-78   Sequence 78, Appl
c 3   36.6   11.9   165156 10 US-10-995-561-13304   Sequence 13304, A
c 4   36.4   11.9   63588   8   US-10-243-735-3   Sequence 3, Appli
  5   36.4   11.9   63588   8   US-10-730-010-3   Sequence 3, Appli
  6   36   11.7   511   4   US-09-925-065A-261418   Sequence 261418,
  7   36   11.7   511   4   US-09-925-065A-261418   Sequence 261418,
  8   36   11.7   548   12 US-10-301-480-339709   Sequence 339709,
  9   36   11.7   548   12 US-10-301-480-953118   Sequence 953118,
 10   36   11.7   661   8   US-09-998-598-965   Sequence 965, App
c11   35.4   11.5   5976   8   US-10-437-963-8075   Sequence 8075, Ap
c12   35.4   11.5   165156 8   US-10-741-601-5668   Sequence 5668, Ap
 13   35.2   11.5   669   12 US-10-301-480-545653   Sequence 545653,
 14   35.2   11.5   669   12 US-10-301-480-1159062   Sequence 1159062,
 15   35   11.4   445   3   US-09-920-300A-620   Sequence 620, App
 16   35   11.4   445   6   US-10-033-528-620   Sequence 620, App
 17   35   11.4   445   7   US-10-099-926-620   Sequence 620, App
 18   35   11.4   445   10 US-10-961-527-620   Sequence 620, App
 19   35   11.4   18820   6   US-09-764-891-6452   Sequence 6452, Ap
 20   35   11.4   18820   3   US-10-705-428-733   Sequence 733, App
 21   34.6   11.3   793   12 US-10-301-480-561365   Sequence 561365,
 22   34.6   11.3   793   12 US-10-301-480-1174774   Sequence 1174774,
 23   34.6   11.3   3673778 7   US-10-312-841-1   Sequence 1, Appli
 24   34.4   11.2   257   8   US-10-424-599-94708   Sequence 94708, A
 25   34.4   11.2   306   8   US-10-085-783A-21758   Sequence 21758, A
 26   34.4   11.2   306   8   US-10-085-783A-21758   Sequence 21758, A
 27   34.4   11.2   321   8   US-10-242-535A-12694   Sequence 12694, A
 28   34.4   11.2   321   8   US-10-085-783A-12694   Sequence 12694, A
 29   34.4   11.2   426   8   US-10-242-535A-39856   Sequence 39856, A
 30   34.4   11.2   426   8   US-10-085-783A-39856   Sequence 39856, A
 31   34.4   11.2   436   8   US-10-242-535A-7614   Sequence 7614, Ap
 32   34.4   11.2   436   8   US-10-085-783A-7614   Sequence 7614, Ap
 33   34.4   11.2   482   8   US-10-242-535A-24125   Sequence 24125, A
 34   34.4   11.2   482   8   US-10-085-783A-24125   Sequence 24125, A
 35   34.4   11.2   533   8   US-10-242-535A-15691   Sequence 15691, A
 36   34.4   11.2   533   8   US-10-085-783A-15691   Sequence 15691, A
 37   34.4   11.2   540   16 US-11-128-061-3720   Sequence 3720, Ap
 38   34.4   11.2   540   16 US-11-128-049-3720   Sequence 3720, Ap
 39   34.4   11.2   541   4   US-09-925-065A-465048   Sequence 465048,
 40   34.4   11.2   541   4   US-09-925-065A-465049   Sequence 465049,
 41   34.4   11.2   541   4   US-09-925-065A-465048   Sequence 465048,
 42   34.4   11.2   541   5   US-09-925-065A-465049   Sequence 465049,
 43   34.4   11.2   544   12 US-10-301-480-521152   Sequence 521152,
 44   34.4   11.2   544   12 US-10-301-480-521153   Sequence 521153,
 45   34.4   11.2   544   12 US-10-301-480-1134561   Sequence 1134561,

                                   ALIGNMENTS

RESULT 1
US-10-651-991-79
; Sequence 79, Application US/10651991
; Publication No. US2005012516lAl
; GENERAL INFORMATION:
;   APPLICANT: CAIRNEY, JOHN
;   APPLICANT: XU, NANFIE
;   TITLE OF INVENTION: DIFFERENTIALLY-EXPRESSED CONIFER cDNAs, AND THEIR USE
;   TITLE OF INVENTION: IN IMPROVING SOMATIC EMBRYOGENSIS
;   FILE REFERENCE: 7648.0023-00
;   CURRENT APPLICATION NUMBER: US/10/651,991
;   CURRENT FILING DATE: 2003-09-02
;   PRIOR APPLICATION NUMBER: 60/239,250
;   PRIOR FILING DATE: 2000-10-11
;   PRIOR APPLICATION NUMBER: 60/260,882
;   PRIOR FILING DATE: 2001-01-12
;   NUMBER OF SEQ ID NOS: 339
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
;   LENGTH: 307
;   TYPE: DNA
;   ORGANISM: Pinus taeda
US-10-651-991-79

  Query Match   100.0%; Score 307; DB 10; Length 307;
  Best Local Similarity   100.0%; Pred. No. 1.6e-81;
  Matches 307; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 GGTGCGATCCAGAATTGCATCGTTGCCGTTGCTACTCCCTTGAATTAGAAGCCTTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GGTGCGATCCAGAATTGCATCGTTGCCGTTGCTACTCCCTTGAATTAGAAGCCTTCC 60

Qy   61 TTGCCACCTGGTCTTGAAGCATCCTATATTAGTCTTATTATTCCTGGCATTATTCCTAGAAT 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTGCCACCTGGTCTTGAAGCATCCTATATTAGTCTTATTATTCCTGGCATTATTCCTAGAAT 120

Qy   121 CCTCTTGCTCCTTAACATCCTATAAGGTCATCTCATGCCAAGAGATCTTGACAAATTTGG 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CCTCTTGCTCCTTAACATCCTATAAGGTCATCTCATGCCAAGAGATCTTGACAAATTTGG 180

```
Qy    181  AAGAATCTGCCAATTGCTCTCCGCCCCTCCTTAACTCGAAGTTGACTCTCTAAGTTGCTGAA  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  AAGAATCTGCCAATTGCTCTCCGCCCCTCCTTAACTCGAAGTTGACTCTCTAAGTTGCTGAA  240

Qy    241  CCTTCTGAGCTACCTGTTCTTTGTTTCTATCATTGCTTTTGTTTCTTTGTGCCATCACTCGGA  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  CCTTCTGAGCTACCTGTTCTTTGTTTCTATCATTGCTTTTGTTTCTTTGTGCCATCACTCGGA  300

Qy    301  TCGCACC  307
           |||||||
Db    301  TCGCACC  307
```

RESULT 2
US-10-651-991-78/c
; Sequence 78, Application US/10651991
; Publication No. US2005012516lA1
; GENERAL INFORMATION:
; APPLICANT: CAIRNEY, JOHN
; APPLICANT: XU, NANFIE
; TITLE OF INVENTION: DIFFERENTIALLY-EXPRESSED CONIFER cDNAs, AND THEIR USE
; TITLE OF INVENTION: IN IMPROVING SOMATIC EMBRYOGENESIS
; FILE REFERENCE: 7648.0023-00
; CURRENT APPLICATION NUMBER: US/10/651,991
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: 60/239,250
; PRIOR FILING DATE: 2000-10-11
; PRIOR APPLICATION NUMBER: 60/260,882
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 339
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 78
; LENGTH: 308
; TYPE: DNA
; ORGANISM: Pinus taeda
US-10-651-991-78

Query Match   96.4%; Score 296; DB 10; Length 308;
Best Local Similarity   99.7%; Pred. No. 3.2e-78;
Matches 307; Conservative   0; Mismatches   0; Indels   1; Gaps   1;

```
Qy      1  GGTGCGATCTAGAATTGCATCTGTTGCCGGTGCT-ACTCCCTTGAATTAGAAGCCTTC   59
           |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    308  GGTGCGATCCTAGAATTGCATCTGTTGCCGGTGCTAACTCCCTTGAATTAGAAGCCTTC  249

Qy     60  CTTGCCACCTGGTCTTGAAGCATCCTTTAGTGTCTTATTATTCCTTGGCATTATTCCTAGAA  119
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248  CTTGCCACCTGGTCTTGAAGCATCCTTTAGTGTCTTATTATTCCTTGGCATTATTCCTAGAA  189

Qy    120  TCCTCTTGCTCCTTAACATCCTATAAGGTCATCTCATGCCAAGAGATCTTGACAAATTTG  179
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    188  TCCTCTTGCTCCTTAACATCCTATAAGGTCATCTCATGCCAAGAGATCTTGACAAATTTG  129

Qy    180  GAAGAATCTGCCAATTGCTCTCCGCCCCTCCTTAACTCGAAGTTGACTCTCTAAGTTGCTGA  239
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    128  GAAGAATCTGCCAATTGCTCTCCGCCCCTCCTTAACTCGAAGTTGACTCTCTAAGTTGCTGA   69

Qy    240  ACCTTCTGAGCTACCTGTTCTTTGTTTCTTTGTTTCATCATTGCTTTTGTTTCTTTGTGCCATCACTCGG  299
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     68  ACCTTCTGAGCTACCTGTTCTTTGTTTCTTTGTTTCATCATTGCTTTTGTTTCTTTGTGCCATCACTCGG    9

Qy    300  ATCGCACC  307
           ||||||||
Db      8  ATCGCACC    1
```

RESULT 3
US-10-995-561-13304/c
; Sequence 13304, Application US/10995561
; Publication No. US2005027205lA1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13304
; LENGTH: 165156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13304

Query Match   11.9%; Score 36.6; DB 10; Length 165156;
Best Local Similarity   47.8%; Pred. No. 32;
Matches 119; Conservative   5; Mismatches 124; Indels   1; Gaps   1;

```
Qy     55  CCTTCTGCCACCTGGTCTTGAAGCATCCTTTAGTCTTATTATTCCTTGGCATTATTCC  114
           ||||||||||| |||  |||  |||||  ||: ::
Db  80842  CCTTCTCCTTCTTTCTTTTGAAGGACCCCTTTTCTCCYTCCTTCCYTCCC  80783

Qy    115  TAGAATCCTCGTGCTCCTTAACATCCTATAAGGTCATCGCCAAGAGATCTTGACAA  174
           |||     ||:  ||||||||:  ||||
Db  80782  TCCCTCCTCCTTCCTTCCTTCCTTTTYTCTTTYTCTTTTCTTTTCTTT  80723

Qy    175  ATTTGGAAGAATCTGCCAATTGCTCCCGCCCCTCCTTAACTC-GAAGTTGACTCTCTAAGT  233
           |||       ||||||||||||:: |||||||
Db  80722  CTTTCCTCCTTCCCCTTCCCCTTCCTTCCTTCCTTCCTCCACCCTCCTCTCCTCTCTCC  80663

Qy    234  TGCTGAACCTTCTGAGCTACCTGTTCTTTGTTTCTATCATTGCTTTTGTTTGTTGCCATC  293
           |||||||| ||||||||
Db  80662  CTTCCTTCCTTCCTTCCYTCCCYCCTTCCTTCCTTCCTTTCTTTTCTTTTCTTTGCCATC  80603

Qy    294  ACTCGGATC  302
           |||||
Db  80602  CACTGTATC  80594
```

RESULT 4
US-10-243-735-3
; Sequence 3, Application US/10243735
; Publication No. US2003022341A1
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001212DIV
; CURRENT APPLICATION NUMBER: US/10/243,735
; CURRENT FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 63588
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(63588)
; OTHER INFORMATION: n = A,T,C or G
US-10-243-735-3

Query Match   11.9%; Score 36.4; DB 6; Length 63588;
Best Local Similarity   51.2%; Pred. No. 24;
Matches   85; Conservative   0; Mismatches   81; Indels   0; Gaps   0;

```
Qy    106  CATTATTCCTAGAATCCTCTTGCTCCTTAACATCCTATAAGGTCATCATGCCAAGAGA  165
           ||||     ||  ||| |||  |||| | |||    |||     |||||||
Db  15012  CAGTCTTACTTGCATGCGCTTTCTCCTCACCCTCCCAGCCTACAGACTTCTTCACAGCAGCACT  15071

Qy    166  TCTTGACAAATTTGGAAGAATCTGCCAATTGCTCCCGCCCCTCCTTAACTCCTTAACTGCTGACT  225
           |||| ||  |||       ||||  ||
Db  15072  TCTCACCCAATGCTGGAGGACTCCCAACTAGACTTGCCCTCCTTAACAGTGAATGGAT  15131
```

```
                    Copyright   GenCore version 5.1.9
                    Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on:      July 11, 2006, 07:46:06 ; Search time 63.2377 Seconds
                                          (without alignments)
                                          6311.756 Million cell updates/sec

Title:       US-09-973-994-79
Perfect score: 307
Sequence:    1 ggtgcgatccagaattgca.........gccatcactcggatcgcacc 307

Scoring table: IDENTITY.NUC
             Gapop 10.0 , Gapext 1.0

Searched:    844069 seqs, 650066433 residues

Total number of hits satisfying chosen parameters:   1688138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    Published_Applications_NA_New:*
         1: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US09_NEW_PUB.seq:*
         2: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US06_NEW_PUB.seq:*
         3: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US07_NEW_PUB.seq:*
         4: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US08_NEW_PUB.seq:*
         5: /EMC_Celerra_SIDS3/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
         6: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US10_NEW_PUB.seq:*
         7: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq:*
         8: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq1:*
         9: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US60_NEW_PUB.seq:*

         Pred. No. is the number of results predicted by chance to have a
         score greater than or equal to the score of the result being printed,
         and is derived by analysis of the total score distribution.

                                   SUMMARIES
```

Result No.		Score	Query % Match	Length	DB	ID	Description
c	1	35.8	11.7	9353	6	US-10-517-441-461	Sequence 461, App
c	2	35	11.4	9353	6	US-10-517-441-108	Sequence 108, App
c	3	34.4	11.2	583	6	US-10-501-187-222	Sequence 222, App
c	4	34.4	11.2	809	8	US-11-266-748A-194452	Sequence 194452,
c	5	34.4	11.2	809	8	US-11-266-748A-227974	Sequence 227974,
c	6	34.4	11.2	961	8	US-11-266-748A-395495	Sequence 395495,
c	7	34.4	11.2	961	8	US-11-266-748A-466541	Sequence 466541,
c	8	34.4	11.2	1033	8	US-11-266-748A-114406	Sequence 114406,
c	9	34.4	11.2	1033	8	US-11-266-748A-120423	Sequence 120423,
c	10	34.4	11.2	1065	8	US-11-266-748A-22072	Sequence 22072, A
c	11	34.2	11.1	9353	6	US-10-517-441-735	Sequence 735, App
c	12	34	11.1	811	8	US-11-266-748A-99299	Sequence 99299, A
c	13	34	11.1	811	8	US-11-266-748A-152110	Sequence 152110,
c	14	33	10.7	1629	8	US-11-217-529-79175	Sequence 79175, A
c	15	32.8	10.7	882	8	US-11-266-748A-227973	Sequence 227973,
c	16	32.6	10.6	737	8	US-11-266-748A-267357	Sequence 267357,
c	17	32.6	10.6	737	8	US-11-266-748A-280601	Sequence 280601,
c	18	32.6	10.6	737	8	US-11-266-748A-327874	Sequence 327874,
c	19	32.6	10.6	873	8	US-11-217-529-5387	Sequence 5387, Ap
c	20	32.4	10.6	1543	7	US-11-218-305-6058	Sequence 6058, Ap
c	21	32.2	10.5	948	8	US-11-266-748A-357437	Sequence 357437,
c	22	32.2	10.5	948	8	US-11-266-748A-440816	Sequence 440816,
c	23	32	10.4	636	8	US-11-266-748A-5648	Sequence 5648, Ap
c	24	32	10.4	636	8	US-11-266-748A-63455	Sequence 63455, A

c	25	32	10.4	636	8	US-11-266-748A-66287	Sequence 66287, A
c	26	32	10.4	2187	7	US-11-218-305-5268	Sequence 5268, Ap
c	27	32	10.4	138443	8	US-11-266-748A-60103	Sequence 60103, A
c	28	31.4	10.2	3120	8	US-11-266-748A-27496	Sequence 27496, A
c	29	31.2	10.2	498	8	US-11-217-529-5270	Sequence 5270, Ap
c	30	31.2	10.2	662	8	US-11-266-748A-43211	Sequence 43211, A
c	31	31.2	10.2	1304	6	US-10-449-902-20764	Sequence 20764, A
c	32	31.2	10.2	4314	8	US-11-217-529-1866	Sequence 1866, Ap
c	33	31	10.1	1689	6	US-10-216-545-4231	Sequence 4231, Ap
c	34	31	10.1	2221	6	US-10-449-902-11700	Sequence 11700, A
c	35	31	10.1	2671	8	US-11-266-748A-182495	Sequence 182495,
c	36	31	10.1	2671	8	US-11-266-748A-243034	Sequence 243034,
c	37	30.8	10.0	488	8	US-11-266-748A-54033	Sequence 54033, A
c	38	30.8	10.0	657	8	US-11-266-748A-74522	Sequence 74522, A
c	39	30.8	10.0	657	8	US-11-266-748A-127333	Sequence 127333,
c	40	30.8	10.0	895	8	US-11-266-748A-52065	Sequence 52065, A
c	41	30.8	10.0	1000	8	US-11-266-748A-195938	Sequence 195938,
c	42	30.8	10.0	1000	8	US-11-266-748A-223473	Sequence 223473,
c	43	30.8	10.0	1000	8	US-11-266-748A-284388	Sequence 284388,
c	44	30.8	10.0	1000	8	US-11-266-748A-291081	Sequence 291081,
c	45	30.8	10.0	1000	8	US-11-266-748A-335817	Sequence 335817,

```
                                         ALIGNMENTS

RESULT 1
US-10-517-441-461/c
; Sequence 461, Application US20060124741
; Publication No. US2006012467A1
; GENERAL INFORMATION:
;  APPLICANT: FOEKENS, John
;  APPLICANT: HARBECK, Nadia
;  APPLICANT: KOENIG, Thomas
;  APPLICANT: MAIER, Sabine
;  APPLICANT: MARTENS, John
;  APPLICANT: MODEL, Fabian
;  APPLICANT: NIMMRICH, Inko
;  APPLICANT: RUJAN, Tamas
;  APPLICANT: SCHMITT, Armin
;  APPLICANT: SCHMITT, Manfred
;  APPLICANT: LOOK, Maxime P.
;  APPLICANT: MARX, Almuth
;  APPLICANT: HOEFLER, Heinz
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast ce
; TITLE OF INVENTION: proliferative disorders
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; CURRENT FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: DE 10317955.0
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: DE 10300036.8
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: DE 10245779.4
; PRIOR FILING DATE: 2002-10-01
; NUMBER OF SEQ ID NOS: 2147
; SEQ ID NO 461
;  LENGTH: 9353
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-517-441-461

 Query Match             11.7%;  Score 35.8;  DB 6;  Length 9353;
 Best Local Similarity   50.3%;  Pred. No. 0.83;
 Matches   88;  Conservative   0;  Mismatches   87;  Indels   0;  Gaps   0;

Qy     19 CATCTGTTGCCGTTGCTACTCCCTTGAATTAGAGAGCCTTCCTTGCCACCTGGTCTTGAA 78
          |||||| ||||| || || ||| || ||| ||| |||| ||| | ||| ||| || |||
Db   4530 CACCATTCCCGCGTCACCCCCACCGTACAATTCTACCACCAACCTAATCTCCTCATCCACAA 4471
```

Qy   79   GCATCCTTTAGTCTTATTATCCTTGGCATTATTCCTAGAATCCTCTTGCTCCTTAACAT 138
        |||| ||||||| ||| ||| || ||||| |||||| ||| |||| ||||||||
Db   4470   AAATACAAATAAAGTTTCACCATACAACTATCTCTAAAATCGTTTTACTTCCTTCATAT 4411

Qy   139   CCTATAAGGTCATCCATGCCAAGAGATCTTGACAAATTGGAAGAAATCTGCCAA 193
        ||| |||
Db   4410   AAATTCAACACATAACTTTTAAATACCTACTCCTATATTTACCAAACACTATAAA 4356

RESULT 2
US-10-517-441-108/c
; Sequence 108, Application US/10517441
; Publication No. US2006012167A1
; GENERAL INFORMATION:
;   APPLICANT: FOEKENS, John
;   APPLICANT: HARBECK, Nadia
;   APPLICANT: KOENIG, Thomas
;   APPLICANT: MAIER, Sabine
;   APPLICANT: MARTENS, John
;   APPLICANT: MODEL, Fabian
;   APPLICANT: NIMMRICH, Inko
;   APPLICANT: RUJAN, Tamas
;   APPLICANT: SCHMITT, Armin
;   APPLICANT: SCHMITT, Manfred
;   APPLICANT: LOOK, Maxime P.
;   APPLICANT: MARX, Almuth
;   APPLICANT: HOEFLER, Heinz
;   TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cel
;   TITLE OF INVENTION: proliferative disorders
;   FILE REFERENCE: 47675-93
;   CURRENT APPLICATION NUMBER: US/10/517,441
;   CURRENT FILING DATE: 2004-12-11
;   PRIOR APPLICATION NUMBER: PCT/EP2003/010881
;   PRIOR FILING DATE: 2003-10-01
;   PRIOR APPLICATION NUMBER: DE 10317955.0
;   PRIOR FILING DATE: 2003-04-17
;   PRIOR APPLICATION NUMBER: DE 10300096.8
;   PRIOR FILING DATE: 2003-01-07
;   PRIOR APPLICATION NUMBER: DE 10245779.4
;   PRIOR FILING DATE: 2002-10-01
;   NUMBER OF SEQ ID NOS: 2147
;   SEQ ID NO 108
;   LENGTH: 9353
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-10-517-441-108

Query Match   11.4%;   Score 35;   DB 6;   Length 9353;
Best Local Similarity   50.3%;   Pred. No. 1.5;
Matches   86;   Conservative   0;   Mismatches   85;   Indels   0;   Gaps   0;

Qy   19   CATCTGTTGCCGTTGCTACTCCCTTTGAATTAGAAGCCTTCCTTGCCACCTGGTCTTGAA 78
        ||| ||||| || |||| || |
Db   4530   CACCATTCCCCGCTGCCCCCCACCGTACAATTCTACACCAACCTGATCTCCTCATCCACAG 4471

Qy   79   GCATCCTTTAGTCTTATTATCCTTGGCATTATTCCTAGAATCCTCTTGCTCCTTAACAT 138
        |||||||
Db   4470   GAATACAAGTGGGCGTTTCACCATACAACTATCTCTAAAGTCGTTTTACTTCCTTCATAT 4411

Qy   139   CCTATAAGGTCATCCATGCCAAGAGATCTTGACAAATTGGAAGAAATCTG 189
        ||| |
Db   4410   GAATTCAACACATAACTTTTGGATACCTACTCCTATATTTGCCAGGCACTG 4360

RESULT 3
US-10-501-187-222/c
; Sequence 222, Application US/10501187
; Publication No. US2006014145SA1
; GENERAL INFORMATION:
;   APPLICANT: Hansen, Rhonda
;   TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED
;   TITLE OF INVENTION: IN CANCEROUS BREAST CELLS AND THEIR METHODS OF USE
;

;   FILE REFERENCE: 2300-17767
;   CURRENT APPLICATION NUMBER: US/10/501,187
;   CURRENT FILING DATE: 2004-07-08
;   PRIOR APPLICATION NUMBER: 60/345,637
;   PRIOR FILING DATE: 2002-01-08
;   NUMBER OF SEQ ID NOS: 516
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 222
;   LENGTH: 583
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-501-187-222

Query Match   11.2%;   Score 34.4;   DB 6;   Length 583;
Best Local Similarity   59.0%;   Pred. No. 0.66;
Matches   59;   Conservative   0;   Mismatches   41;   Indels   0;   Gaps   0;

Qy   68   CTGGTCTTGAAGCATCCTTTAGTCTTATTATCCTTGGCATTATTCCTAGAATCCTCTTG 127
        | || |||||| |||
Db   324   CAGGTTTTTTGGCTTTCTTTGATTCATATTGTGCAAGACGTTCTTCCCTTAGCCTCTTTG 265

Qy   128   CTCCTTAACATCCTATAAGGTCATCCATGCCAAGAGATC 167
        ||| ||| ||| |||||
Db   264   CTTCTTCACTTTCCTCCTCATCATCAGATCCAAGAGGGTC 225

RESULT 4
US-11-266-748A-194452/c
; Sequence 194452, Application US/11266748A
; Publication No. US2006013463A1
; GENERAL INFORMATION:
;   APPLICANT: Harkin, Paul
;   APPLICANT: Johnston, Patrick
;   APPLICANT: Mulligan, Karl
;   TITLE OF INVENTION: Transcriptome Microarray Technology and
;   TITLE OF INVENTION: Methods of Using the Same
;   FILE REFERENCE: 55815-0102 (319189)
;   CURRENT APPLICATION NUMBER: US/11/266,748A
;   CURRENT FILING DATE: 2005-11-03
;   PRIOR APPLICATION NUMBER: EP 04105479.2
;   PRIOR FILING DATE: 2004-11-03
;   PRIOR APPLICATION NUMBER: EP 04105482.6
;   PRIOR FILING DATE: 2004-11-03
;   PRIOR APPLICATION NUMBER: EP 04105483.4
;   PRIOR FILING DATE: 2004-11-03
;   PRIOR APPLICATION NUMBER: EP 04105507.0
;   PRIOR FILING DATE: 2004-11-03
;   PRIOR APPLICATION NUMBER: EP 04105485.9
;   PRIOR FILING DATE: 2004-11-03
;   PRIOR APPLICATION NUMBER: EP 04105484.2
;   PRIOR FILING DATE: 2004-11-03
;   PRIOR APPLICATION NUMBER: US 60/662,276
;   PRIOR FILING DATE: 2005-03-14
;   PRIOR APPLICATION NUMBER: US 60/700,293
;   PRIOR FILING DATE: 2005-07-18
;   NUMBER OF SEQ ID NOS: 483996
;   SOFTWARE: Patentin version 3.3
;   SEQ ID NO 194452
;   LENGTH: 809
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-194452

Query Match   11.2%;   Score 34.4;   DB 8;   Length 809;
Best Local Similarity   59.0%;   Pred. No. 0.76;
Matches   59;   Conservative   0;   Mismatches   41;   Indels   0;   Gaps   0;

Qy   68   CTGGTCTTGAAGCATCCTTTAGTCTTATTATCCTTGGCATTATTCCTAGAATCCTCTTG 127
        | || |||||| |||
Db   490   CAGGTTTTTTGGCTTTCTTTGATTCATATTGTGCAAGACGTTCTTCCCTTAGCCTCTTTG 431

Qy   128   CTCCTTAACATCCTATAAGGTCATCCATGCCAAGAGATC 167
        ||| ||| ||| |||||

```
          GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.
```

OM nucleic - nucleic search, using sw model

Run on: July 11, 2006, 07:01:45 ; Search time 392.398 Seconds
                                     (without alignments)
                                     5454.872 Million cell updates/sec

Title: US-09-973-994-79
Perfect score: 307
Sequence: 1 ggtgcgatcctagaattgca.........gcatcactcggatcgcacc 307

Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database : N Geneseq_8:*
            1: geneseqn1980s:*
            2: geneseqn1990s:*
            3: geneseqn2000s:*
            4: geneseqn2001as:*
            5: geneseqn2001bs:*
            6: geneseqn2002as:*
            7: geneseqn2002bs:*
            8: geneseqn2003as:*
            9: geneseqn2003bs:*
           10: geneseqn2003cs:*
           11: geneseqn2003ds:*
           12: geneseqn2004as:*
           13: geneseqn2004bs:*
           14: geneseqn2005s:*
           15: geneseqn2006s:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	307	100.0	307	14	AEA49584	Aea49584 Loblolly
c	2	296	96.4	308	14	AEA49583	Aea49583 Loblolly
	3	36.4	11.9	63588	8	ABS57150	Abs57150 Human gen
c	4	36.2	11.8	2000	11	ACL36306	Acl36306 Rice stre
c	5	36	11.7	661	6	ABV87654	Abv87654 Human col
c	6	35.8	11.7	677	1	AAN60393	Aan60393 Sequence
c	7	35.8	11.7	9353	13	ADS89445	Ads89445 Oligonucl
c	8	35.6	11.6	265118	5	AAH41227	Aah41227 Pyrococcu
c	9	35.4	11.5	165156	13	ADS36459	Ads36459 Human aut
	10	35	11.4	445	6	ABK45069	Abk45069 cDNA enco
	11	35	11.4	2607	14	ADW16473	Adw16473 Eucalyptu
c	12	35	11.4	9353	13	ADS89092	Ads89092 Human AKR
	13	35	11.4	18820	4	AAL03764	Aal03764 Human rep
	14	35	11.4	18820	4	ABA07938	Aba07938 Human ova
c	15	34.8	11.3	64976	2	AAV21209_16	Continuation (17 o
c	16	34.6	11.3	5394	2	AAT00872	Aat00872 Murine mC
c	17	34.6	11.3	5394	14	AAZ86916	Aaz86916 Mouse mC2
c	18	34.6	11.3	5394	14	AEC05449	Aec05449 Mouse mC2
c	19	34.4	11.2	521	10	AB284664	Abz84664 Toxicolog
c	20	34.4	11.2	540	14	AEE14209	Aee14209 Hamster S
c	21	34.4	11.2	553	6	ABV89065	Abv89065 Human col
	22	34.4	11.2	561	6	ABK44654	Abk44654 cDNA enco
	23	34.4	11.2	563	6	ABK45341	Abk45341 cDNA enco
c	24	34.4	11.2	579	14	AEE10567	Aee10567 Hamster c
c	25	34.4	11.2	583	10	ADK11816	Adk11816 Breast ca
c	26	34.4	11.2	583	14	ACL53983	Acl53983 Human col
	27	34.4	11.2	619	6	ABK54823	Abk54823 Human col
	28	34.4	11.2	741	6	ABQ55781	Abq55781 Human ova
c	29	34.4	11.2	818	10	ADG32718	Adg32718 Human DNA
c	30	34.4	11.2	821	13	ACN37817	Acn37817 Tumour-as
	31	34.4	11.2	821	14	ADZ48889	Adz48889 Insulin s
c	32	34.4	11.2	961	13	ACN37818	Acn37818 Tumour-as
	33	34.4	11.2	961	14	ADZ70690	Adz70690 Human cDN
c	34	34.4	11.2	1089	4	AAK53287	Aak53287 Human pol
c	35	34.4	11.2	1126	4	AAK52303	Aak52303 Human pol
c	36	34.4	11.2	1127	12	ADI13373	Adi13373 Human ste
	37	34.4	11.2	1147	4	AAD08441	Aad08441 Human sec
	38	34.4	11.2	1147	5	AAD05439	Aad05439 Human sec
c	39	34.4	11.2	1155	6	AAS85654	Aas85654 DNA encod
c	40	34.4	11.2	1343	14	AED73250	Aed73250 Human pla
c	41	34.4	11.2	2453	13	ADX54786	Adx54786 plant ful
	42	34.2	11.2	80963	14	AEA60_095_3	AEA60_095_3 Continuation (4 of
c	43	34	11.1	9353	13	ADS89719	Ads89719 Oligonucl
c	44	34	11.1	562	3	AAF16377	Aaf16377 Human pro
c	45	34	11.1	659	6	ABQ60765	Abq60765 Human col

ALIGNMENTS

```
RESULT 1
AEA49584
ID   AEA49584 standard; cDNA; 307 BP.
XX
AC   AEA49584;
XX
DT   11-AUG-2005  (first entry)
XX
DE   Loblolly pine cDNA clone, LPS-096; SEQ ID 79.
XX
KW   Plant; embryogenesis; plant breeding; paper; wood; ss.
XX
OS   Pinus taeda.
XX
PN   US2005125161-A1.
XX
PD   09-JUN-2005.
XX
PF   02-SEP-2003; 2003US-00651991.
XX
PR   11-OCT-2000; 2000US-0239250P.
PR   12-JAN-2001; 2001US-0260882P.
PR   11-OCT-2001; 2001US-00973994.
XX
PA   (PAPE-) INST PAPER SCI & TECHNOLOGY INC.
XX
PI   Cairney J, Xu N;
XX
DR   WPI; 2005-417092/42.
XX
PT   Relational database of cDNA molecules including those corresponding to
PT   Loblolly pine major intrinsic protein, being differentially expressed
PT   during plant embryogenesis, useful for staging plant embryos.
XX
PS   Claim 17; SEQ ID NO 79; 219pp; English.
XX
CC   The present invention relates to a relational database of cDNA molecules
CC   comprising multiple nucleotide sequences (AEA49506-AEA49832). The
CC   database is useful for staging plant embryos, selecting advantageous
CC   plant clones, determining embryo fitness, and selecting advantageous
CC   conditions for embryo development. The invention is particularly intended
```

CC   for use in breeding trees with superior characteristics for use in
CC   production of paper and other wood products.
XX
SQ   Sequence 307 BP; 62 A; 84 C; 51 G; 110 T; 0 U; 0 Other;

Query Match   100.0%; Score 307; DB 14; Length 307;
Best Local Similarity   100.0%; Pred. No. 1.2e-79;
Matches 307; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 GGTGCGATCCTAGAATTGCATCGTGTGCCGGTTGCTACTCCCTTGAATTAGAAGCCTTCC 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 GGTGCGATCCTAGAATTGCATCGTGTGCCGGTTGCTACTCCCTTGAATTAGAAGCCTTCC 60

Qy   61 TTGCCACCTGGTCTTGAAGCATCCTTTAGTCTTATTATCCTTGGCATTATTCCTAGAAT 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTGCCACCTGGTCTTGAAGCATCCTTTAGTCTTATTATCCTTGGCATTATTCCTAGAAT 120

Qy  121 CCTCTTGCTCCTTAACATCCTATAAGGTCATCATGCCAAGAGATCTTGACAAATTGG 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CCTCTTGCTCCTTAACATCCTATAAGGTCATCATGCCAAGAGATCTTGACAAATTGG 180

Qy  181 AAGAATCTGCCAATTGCTCCCGGCCCTCCTTAACTCGAAGTTGACTCTCCAAGTTGCTGAA 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 AAGAATCTGCCAATTGCTCCCGGCCCTCCTTAACTCGAAGTTGACTCTCCAAGTTGCTGAA 240

Qy  241 CCTTCTGAGCTACCTGTTCTTTGTTTCTATCATTGCTTTCTTTGTGCCATCACTCGGA 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 CCTTCTGAGCTACCTGTTCTTTGTTTCTATCATTGCTTTCTTTGTGCCATCACTCGGA 300

Qy  301 TCGCACC 307
        |||||||
Db  301 TCGCACC 307
```

RESULT 2
AEA49583/c
ID   AEA49583 standard; cDNA; 308 BP.
XX
AC   AEA49583;
XX
DT   11-AUG-2005 (first entry)
XX
DE   Loblolly pine cDNA clone, LPS-095, SEQ ID 78.
XX
KW   Plant; embryogenesis; plant breeding; paper; wood; ss.
XX
OS   Pinus taeda.
XX
PN   US2005125161-A1.
XX
PD   09-JUN-2005.
XX
PF   02-SEP-2003; 2003US-00651991.
XX
PR   11-OCT-2000; 2000US-0239250P.
PR   12-JAN-2001; 2001US-0260882P.
PR   11-OCT-2001; 2001US-00973994.
XX
PA   (PAPE-) INST PAPER SCI & TECHNOLOGY INC.
XX
PI   Cairney J, Xu N;.
XX
DR   WPI; 2005-417092/42.
XX
PT   Relational database of cDNA molecules including those corresponding to
PT   Loblolly pine major intrinsic protein, being differentially expressed
PT   during plant embryogenesis, useful for staging plant embryos.
XX
PS   Claim 17; SEQ ID NO 78; 219pp; English.
XX
CC   The present invention relates to a relational database of cDNA molecules
CC   comprising multiple nucleotide sequences (AEA49506-AEA49832). The
CC   database is useful for staging plant embryos, selecting advantageous
CC   plant clones, determining embryo fitness, and selecting advantageous
CC   conditions for embryo development. The invention is particularly intended
CC   for use in breeding trees with superior characteristics for use in
CC   production of paper and other wood products.
XX
SQ   Sequence 308 BP; 110 A; 51 C; 84 G; 63 T; 0 U; 0 Other;

Query Match   96.4%; Score 296; DB 14; Length 308;
Best Local Similarity   99.7%; Pred. No. 2e-76;
Matches 307; Conservative 0; Mismatches 0; Indels 0; Gaps 1;

```
Qy    1 GGTGCGATCCTAGAATTGCATCGTGTGCCGGTTGCT-ACTCCCTTGAATTAGAAGCCTTC 59
        ||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db  308 GGTGCGATCCTAGAATTGCATCGTGTGCCGGTTGCTAACTCCCTTGAATTAGAAGCCTTC 249

Qy   60 CTTGCCACCTGGTCTTGAAGCATCCTTTAGTCTTATTATCCTTGGCATTATTCCTAGAA 119
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  248 CTTGCCACCTGGTCTTGAAGCATCCTTTAGTCTTATTATCCTTGGCATTATTCCTAGAA 189

Qy  120 TCCTCTTGCTCCTTAACATCCTATAAGGTCATCATGCCAAGAGATCTTGACAAATTGG 179
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  188 TCCTCTTGCTCCTTAACATCCTATAAGGTCATCATGCCAAGAGATCTTGACAAATTG 129

Qy  180 GAAGAATCTGCCAATTGCTCCCGGCCCTCCTTAACTCGAAGTTGACTCTCCAAGTTGCTGA 239
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  128 GAAGAATCTGCCAATTGCTCCCGGCCCTCCTTAACTCGAAGTTGACTCTCCAAGTTGCTGA 69

Qy  240 ACCTTCTGAGCTACCTGTTCTTTGTTTCTATCATTGCTTTCTTTGTGCCATCACTCGG 299
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   68 ACCTTCTGAGCTACCTGTTCTTTGTTTCTATCATTGCTTTCTTTGTGCCATCACTCGG 9

Qy  300 ATCGCACC 307
        ||||||||
Db    8 ATCGCACC 1
```

RESULT 3
ABS57150
ID   ABS57150 standard; DNA; 63588 BP.
XX
AC   ABS57150;
XX
DT   04-FEB-2003 (first entry)
XX
DE   Human gene encoding a serine/threonine kinase.
XX
KW   Human; ds; gene; serine/threonine kinase; lung carcinoma; chromosome 13;
KW   SNP; single nucleotide polymorphism.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   variation   replace(1848,G)
FT   /*tag= am
FT   /standard_name= "Single nucleotide polymorphism"
FT   variation   replace(1593,C)
FT   /*tag= an
FT   /standard_name= "Single nucleotide polymorphism"
FT   CDS   2012..60590
FT   /*tag= a
FT   /product= "Ser/Thr kinase"
FT   exon   2012..2128
FT   /*tag= b
FT   /number= 1
FT   intron   2129..10078
FT   /*tag= c
FT   /number= 1
FT   variation   replace(3796,T)
FT   /*tag= ao
FT   /standard_name= "Single nucleotide polymorphism"
FT   variation   replace(4854,G)
FT   /*tag= ap
FT   /standard_name= "Single nucleotide polymorphism"

OM nucleic - nucleic search, using sw model

```
          GenCore version 5.1.9
          Copyright (c) 1993 - 2006 Bioceleration Ltd.
```

Run on: July 11, 2006, 07:10:56 ; Search time 3077.57 Seconds
                                    (without alignments)
                                    5578.184 Million cell updates/sec

Title: US-09-973-994-79
Perfect score: 307
Sequence: 1 ggtgcgatccagaattgca.........gccatcactcgatcgcacc 307
Scoring table: IDENTITY_NUC
          Gapop 10.0 , Gapext 1.0

Searched: 4836798 seqs, 2795965780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database : EST:*
          1: gb_est1:*
          2: gb_est3:*
          3: gb_est4:*
          4: gb_est5:*
          5: gb_est6:*
          6: gb_htc:*
          7: gb_est2:*
          8: gb_est7:*
          9: gb_est8:*
         10: gb_gss1:*
         11: gb_gss1:*
         12: gb_gss2:*
         13: gb_gss3:*
         14: gb_gss4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES
```

| Result No. | Score | % Match | Query Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 44.2 | 14.4 | 482 | 12 | BZ833608 | BZ833608 CH240 258 |
| 2 | 42.6 | 13.9 | 997 | 14 | CNS005TE | AL060767 Drosophil |
| 3 | 42.4 | 13.8 | 466 | 1 | AL794888 | AL794888 AL794888 |
| 4 | 42.4 | 13.8 | 862 | 10 | DT441734 | DT441734 JGI_CABK3 |
| 5 | 41.4 | 13.5 | 738 | 14 | BX159004 | BX159004 DanIo rer |
| 6 | 40.8 | 13.3 | 656 | 1 | AL896255 | AL896255 AL896255 |
| c 7 | 40 | 13.0 | 911 | 14 | AG887137 | AG887137 Oryza sat |
| c 8 | 40 | 13.0 | 936 | 11 | AZ209923 | AZ209923 SP 0150 B |
| c 9 | 39.2 | 12.7 | 694 | 10 | DT853440 | DT853440 LB0058.CR |
| c 10 | 39 | 12.7 | 791 | 10 | DV553586 | DV553586 DC05004D0 |
| c 11 | 38.8 | 12.6 | 532 | 4 | CD056130 | CD056130 HO11B04S |
| c 12 | 38.8 | 12.6 | 617 | 8 | CO010537 | CO010537 EST798872 |
| c 13 | 38.6 | 12.6 | 763 | 5 | CJ460131 | CJ460131 CJ460131 |
| c 14 | 38.6 | 12.6 | 808 | 9 | CX444624 | CX444624 JGI_XZG10 |
| 15 | 38.6 | 12.6 | 939 | 14 | DU958506 | DU958506 225686 To |
| 16 | 38.4 | 12.5 | 549 | 10 | W52992 | W52992 zc02e10.r1 |
| 17 | 38.4 | 12.5 | 595 | 10 | DR431991 | DR431991 nax50dil. |
| 18 | 38.4 | 12.5 | 653 | 2 | BJ636126 | BJ636126 BJ636126 |
| 19 | 38.4 | 12.5 | 800 | 2 | BJ643839 | BJ643839 BJ643839 |
| 20 | 38.4 | 12.5 | 1017 | 11 | | AQ325859 nbxb0021H |
| 21 | 38 | 12.4 | 728 | 13 | | CW262050 104 729 1 |
| c 22 | 38 | 12.4 | 753 | 10 | | BH452101 BOHNU61TR |
| c 23 | 38 | 12.4 | 1101 | 10 | | DV012710 CNB276-C0 |
| c 24 | 37.8 | 12.3 | 440 | 2 | | BI449723 dae74g01. |
| 25 | 37.8 | 12.3 | 442 | 10 | | DT467317 GH_CHX20D |
| 26 | 37.8 | 12.3 | 791 | 10 | | DT465391 GH_CHX17E |
| c 27 | 37.8 | 12.2 | 1101 | 14 | | AL107216 Drosophil |
| 28 | 37.6 | 12.2 | 485 | 9 | | AA782790 aj08g07.8 |
| c 29 | 37.6 | 12.2 | 744 | 9 | | CX458263 JGI_XZG27 |
| c 30 | 37.6 | 12.2 | 765 | 9 | | CX413015 JGI_XZT29 |
| c 31 | 37.6 | 12.2 | 905 | 13 | | BF665751 60I24084 |
| c 32 | 37.4 | 12.2 | 496 | 13 | | CL894723 abg27h05. |
| c 33 | 37.4 | 12.2 | 850 | 10 | | DR898269 JGI_XZT45 |
| c 34 | 37.4 | 12.2 | 1101 | 14 | | AL108811 Drosophil |
| 35 | 37.4 | 12.2 | 1298 | 14 | | AG311035 Mus muscu |
| 36 | 37 | 12.1 | 814 | 14 | | BX995750 Reverse s |
| 37 | 37 | 12.1 | 893 | 2 | | AL553628 AL553628 |
| c 38 | 37 | 12.1 | 1213 | 2 | | BI560283 603255520 |
| c 39 | 36.8 | 12.0 | 835 | 11 | | BZ251414 CH230-262 |
| 40 | 36.8 | 12.0 | 901 | 13 | | C2984248 19899I To |
| 41 | 36.8 | 12.0 | 949 | 13 | | DU058624 90441 Tom |
| c 42 | 36.6 | 11.9 | 763 | 14 | | AL269608 Tetraodon |
| c 43 | 36.6 | 11.9 | 785 | 5 | | CP285981 AGENCOURT |
| 44 | 36.6 | 11.9 | 866 | 10 | | DV036164 BRS1655 s |
| 45 | 36.4 | 11.9 | 715 | 11 | | BH297774 CH230-57H |

```
                                ALIGNMENTS

RESULT 1
BZ833608
LOCUS       BZ833608              482 bp    DNA    linear    GSS 18-MAR-2003
DEFINITION  CH240_258K11.TV CHORI-240 Bos taurus genomic clone CH240_258K11,
            genomic survey sequence.
ACCESSION   BZ833608
VERSION     BZ833608.1  GI:29060400
KEYWORDS    GSS.
SOURCE      Bos taurus (cattle)
ORGANISM    Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 482)
AUTHORS     Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
            Shvartsbeyn,A., Gebregeorgis,E., Chen,D., Riggs,F., de Jong,P.,
            Crawford,A.M. and McEwan,J.C.
TITLE       Bovine BAC End Sequences from Library CHORI-240
JOURNAL     Unpublished (2003)
COMMENT     Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the bovine BAC library CHORI-240
            (http://www.chori.org/bacpac/bovine240.htm). For BAC library
            availability, please contact Pieter de Jong (pdejong@mail.cho.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/ordering information.htm). This work
            was undertaken as part of the International Bovine BAC Mapping
            Consortium (IBBMC) by AgResearch Ltd., New Zealand and The
            Institute of Genomic Research (TIGR), USA.
            Plate: 258 row: K column: 11
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..482
                     /organism="Bos taurus"
                     /mol_type="genomic DNA"
                     /strain="breed: Hereford"
```

```
                            /db_xref="taxon:9913"
                            /clone="CH240_258K11"
                            /sex="Male"
                            /cell_type="Blood"
                            /clone_lib="CHORI-240"
                            /note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
                            Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
                            library (Male) produced by Pieter de Jong"

ORIGIN

      Query Match          14.4%;   Score 44.2;  DB 12;  Length 482;
      Best Local Similarity 51.9%;  Pred. No. 0.28;
      Matches 123;  Conservative  0;  Mismatches 113;  Indels 1;  Gaps 1;

Qy   51 GAAGCCTTCCTTGCCACCTGGTCTTGAAGCATCCTTTAGTCTTATTATCCTTGCATTA  110
            ||||  |||||||  ||||||||||||||||||||||||||||||||||||
Db   19 GATGTTTTTCTATTCCTGGTTTAAGAGTATTGGGGTTTTTGTTAAAGGCTGTTTCTATA   78

Qy  111 TTCCTAGAATCCTCCTTGCTCCTTAACATCCTATAAGGTCATCTCATGCCAAGAGATCTTG 170
            ||||||||||||||||||  ||| || | ||| |||||
Db   79 TTGATATAATCATAATTCCCTCTTTAGCATATATTCTTTTAAATTGAAAAAATTTTA    138

Qy  171 ACAAATTGGAAGAATCTGCCAATTGCTCCCGCCCTCCTTAACTCGAAGTTGACTCCTCTA  230
            |||||||   |||| ||||
Db  139 ATGATCTCACTTATATGTGGAATTGCTCACTTTTAAAATACTCAATCTTATGTTTGTA   198

Qy  231 AGTTGCTGAACCTTCTG-AGCTACCTGTTCTTGTTTTCTATCATTGCTTTTCTTTG     286
            || ||||||||| ||| |||||||||| ||  ||  ||  || |||
Db  199 GATTCATGAGTCTTCTGCAGTGATTGTGCTTTGATTGTGTATATTTTCACTGCTGTACTCTG 255


RESULT 2
CNS005TE
LOCUS       CNS005TE        997 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR12K22 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL060767
VERSION     AL060767.1  GI:4943573
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
ORGANISM    Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 997)
AUTHORS     Genoscope.
TITLE       Direct Submission
JOURNAL     Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr).
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES    Location/Qualifiers
source      1..997
            /organism="Drosophila melanogaster"
            /mol_type="genomic DNA"
            /db_xref="taxon:7227"
            /clone="BACR12K22"
            /clone_lib="RPCI-98"
```

```
                            /note="end : TET3"

ORIGIN

      Query Match          13.9%;   Score 42.6;  DB 14;  Length 997;
      Best Local Similarity 18.1%;  Pred. No. 0.94;
      Matches 48;  Conservative 104;  Mismatches 113;  Indels 0;  Gaps 0;

Qy   21 TCTGTTGCCGTTGCTACTCCCTTGAATTAGAAGCCTTCCTTGCCACCTGGTCTTGAAGC   80
           :: || |:: ||  |: :::::   :::::::::  :::: ::: :: ::: :: : :
Db  478 YYTTYTTYTTYTYYYYTCTTYYCYYCCYYTYCCYYYYYYTYTYYTYTYCYCYTYTCTY    537

Qy   81 ATCCTTTAGTCTATTATCCTTGGCATTATCCTTGCTCCTTAACATCC              140
           ||::: :: :::   ::::::::: : :::::::::: :::: :::
Db  538 CTCYYTTYYYYYTYYTYYYYTYCYYYYYYYYYYYCYYTCYYCYYYYYCY            597

Qy  141 TATAAGGTCATCTCATGCCAAGAGATCTTGACAAATTGGAAGAATCTGCCAATTGCTCC  200
           ::: ::::::::::::::  ::::  ::::::::::::::::: ::::::::
Db  598 YTCYYYYCYYYCYCYYYYCYYYYYYYCYYYYYCYYYYYYYY                    657

Qy  201 CGCCCTCCTTAACTCGAAGTTGACTCCTCAAGTTGCTGAACCTTCTGAGCTACCTGTTCT 260
           :: :::: ::::::::  :: :::::::::::::::::::::::::  ::: : :::::
Db  658 AYKYCYYYTCTCYYCYCYCYYYYTCTCYCYYYYTTYTYYTYYTYTYT              717

Qy  261 TTGTTTCTATCATTGCTTTTCTTTG                                    285
           ::: :: ::::: ::::::::::::
Db  718 CTCYYYYYYYYYYYYTTTTTCTYYT                                    742


RESULT 3
AL794888/c
LOCUS       AL794888        466 bp    mRNA    linear   EST 13-NOV-2003
DEFINITION  AL794888 XGC-neurula Xenopus tropicalis cDNA clone TNeu15b01 5',
            mRNA sequence.
ACCESSION   AL794888
VERSION     AL794888.2  GI:38316656
KEYWORDS    EST.
SOURCE      Xenopus tropicalis (western clawed frog)
ORGANISM    Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1 (bases 1 to 466)
AUTHORS     Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
TITLE       Sanger Xenopus tropicalis EST project 2001 (11_2003)
JOURNAL     Unpublished (2003)
COMMENT     On Jun 25, 2002 this sequence version replaced gi:21580592.
            Contact: Taylor R
            Sanger Institute
            Hinxton, Cambridgeshire, CB10 1SA, UK
            Email: trop@sanger.ac.uk
            This sequence is from a Xenopus Gene Collection (XGC) library
            constructed by Aaron M. Zorn.
            cDNA was oligo dT primed from 5ug of poly A+ RNA from neurula.
            EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
            5' end and NotI at the 3' end.
            Vector: pCS107; Site 1: EcoRI; Site_2: NotI
            Host: Escherichia coli DH10B
            Sanger Xenopus tropicalis EST project 2001
            TROPICALIS_SEQUENCE_ID: TNeu15b01.p1cSP6
            Sequencing_primer: SP6.
FEATURES    Location/Qualifiers
source      1..466
            /organism="Xenopus tropicalis"
            /mol_type="mRNA"
            /db_xref="taxon:8364"
            /clone="TNeu15b01"
            /dev_stage="neurula"
            /lab_host="Escherichia coli DH10B"
            /clone_lib="XGC-neurula"
            /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
            was oligo dT primed from 5ug of poly A+ RNA from neurula.
            EcoRI-NotI cut cDNA was then ligated into pCS107 with
            EcoRI at the 5' end and NotI at the 3' end."
```

```
          Copyright  GenCore version 5.1.9
          Copyright (c) 1993 - 2006  Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on:   July 11, 2006, 07:02:26 ; Search time 2092.79 Seconds
                        (without alignments)
                        9380.717 Million cell updates/sec

Title:    US-09-973-994-79
Perfect score: 307
Sequence:  1 ggtgcgatcctagaattgca........gccatcactcggatcgcacc 307

Scoring table:  IDENTITY.NUC
                Gapop 10.0 , Gapext 1.0

Searched:   636136 segs, 3197371052S residues

Total number of hits satisfying chosen parameters:  12732272

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
             1: gb_env:*
             2: gb_pat:*
             3: gb_ph:*
             4: gb_pl:*
             5: gb_pr:*
             6: gb_ro:*
             7: gb_sts:*
             8: gb_sy:*
             9: gb_un:*
            10: gb_vi:*
            11: gb_ov:*
            12: gb_htg:*
            13: gb_in:*
            14: gb_om:*
            15: gb_ba:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                    SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description
1	55.8	18.2	7218	2	I66494	I66494 Sequence 14
2	42.6	13.9	230556	12	AC162231	AC162231 Bos tauru
3	42	13.7	110000	12	AC121678_4	Continuation (5 of
4	40.4	13.2	230071	12	AC132699	AC132699 Rattus no
5	40.4	13.2	338915	12	AC130523	AC130523 Rattus no
c 6	39.8	13.0	45027	13	CBRG39N14	AC084575 Caenorhab
c 7	39	12.7	224084	12	CR318629	CR318629 Danio rer
8	39	12.7	237446	11	BX546500	BX546500 Zebrafish
9	38.8	12.6	110000	12	AC107427_1	Continuation (2 of
10	38.6	12.6	7761	10	OPU83594	U83594 Ovine papil
11	38.4	12.5	171562	5	AC107027	AC107027 Homo sapi
c 12	38.4	12.5	197206	12	AC023757	AC023757 Homo sapi
c 13	38.4	12.5	222975	12	AC130905	AC130905 Rattus no
14	38.4	12.5	234072	12	AC099135	AC099135 Rattus no
15	38.2	12.4	183481	12	AC120912	AC120912 Rattus no
16	38.2	12.4	234844	12	AC111734	AC111734 Rattus no
c 17	38	12.4	187571	12	AC177302	AC177302 Strongylo
18	38	12.4	221478	12	AC079545	AC079545 Mus muscu
19	37.8	12.3	110000	4	CR382130_32	Continuation (33 o
c 20	37.8	12.3	116409	6	AC131118	AC131118 Mus muscu
21	37.8	12.3	153310	13	AC159433	AC159433 Trypanoso
22	37.8	12.3	155548	11	AC159413	AC159413 Trypanoso
23	37.8	12.3	206480	11	BX255911	BX255911 Zebrafish
c 24	37.6	12.2	184463	6	AC126553	AC126553 Mus muscu
c 25	37.6	12.2	216501	6	AC107795	AC107795 Mus muscu
26	37.6	12.2	219784	6	AC112792	AC112792 Mus muscu
c 27	37.4	12.2	156045	12	AC144885	AC144885 Papio anu
c 28	37.4	12.2	165583	12	AC119613	AC119613 Rattus no
c 29	37.4	12.2	166616	12	AC145761	AC145761 Papio anu
c 30	37.4	12.2	201602	12	AC074146	AC074146 Mus muscu
c 31	37.4	12.2	237286	12	AC113645	AC113645 Rattus no
32	37.2	12.1	110000	4	AP008218_088	Continuation (89 o
c 33	37.2	12.1	151176	4	CNS08CCS	AL954154 Oryza sat
34	37.2	12.1	348034	13	CR382400	CR382400 Plasmodiu
c 35	37	12.1	52052	12	AC100081	AC100081 Mus muscu
c 36	37	12.1	123748	6	BX678774	BX678774 Mouse DNA
c 37	37	12.1	177857	12	AC135432	AC135432 Rattus no
c 38	37	12.1	202838	12	CR388230	CR388230 Mus muscu
c 39	36.8	12.0	85903	4	AC177439	AC177439 Strongylo
40	36.8	12.0	120077	4	AC148292	AC148292 Medicago
41	36.8	12.0	137281	12	AC124217	AC124217 Medicago
42	36.8	12.0	149635	11	BX510957	BX510957 Zebrafish
43	36.8	12.0	217726	6	AL928680	AL928680 Mouse DNA
c 44	36.8	12.0	254499	12	AC133758	AC133758 Rattus no
45	36.6	11.9	70932	12	AC174554	AC174554 Strongylo

```
                    ALIGNMENTS

RESULT 1
I66494
LOCUS       I66494         7218 bp    DNA     linear   PAT 28-DEC-1997
DEFINITION  Sequence 14 from patent US5670367.
ACCESSION   I66494
VERSION     I66494.1  GI:2724471
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 7218)
  AUTHORS   Dorner,F., Scheiflinger,F. and Falkner,F.Gunter.
  TITLE     Recombinant fowlpox virus.
  JOURNAL   Patent: US 5670367-A 14 23-SEP-1997;
FEATURES             Location/Qualifiers
     source          1..7218
                     /organism="unknown"
                     /mol_type="unassigned DNA"

ORIGIN

Query Match             18.2%;  Score 55.8;  DB 2;  Length 7218;
Best Local Similarity   2.0%;  Pred. No. 4.7e-07;
Matches    6;  Conservative 186; Mismatches 103; Indels 0;  Gaps 0;

Qy      3  TGCGATCCTAGAATTGCATCGTTGCCGGTTGCTACTCCCCTTGAATTAGAAGCCTTCCTT 62
           ||||||:::                                          ::|||::
Db   1061  TGCGGATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1120

Qy     63  GCCACCTGGTCTTGAAGCATCCTTTAGTCTATTATTCCTTGGCATTATTCCTAGAATCC 122
           :: ::::         :::::::
Db   1121  YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1180

Qy    123  TCTTGCTCCTTAACATCCTATAAGGCATCTCATGCCAAGAGATCTTGACAAATTGGAA 182
           ::: :::: :::: :::::
Db   1181  YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1240

Qy    183  GAATCTGCCAATTGCTCCCGCCCTCCTTAACTGAAGTCTCTCTAAGTTGCTGAACC 242
           :: ::: :: ::
Db   1241  YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1300

Qy    243  TTCTGAGCTACCCTGTTCTTTGTTTTCTATCATTGCTTTCTTTGTGCCATCACTC 297
```

```
Db          1301 :::: :: :::: :: :::::: :: :::::::: :: ::::::::: :: ::::::::::
                 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1355

RESULT 2
AC162231
LOCUS       AC162231              230556 bp    DNA     linear   HTG 01-JUL-2005
DEFINITION  Bos taurus clone CH240-112A3, WORKING DRAFT SEQUENCE, 24 unordered
            pieces.
ACCESSION   AC162231
VERSION     AC162231.2  GI:68302821
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Bos taurus (cattle)
ORGANISM    Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 230556)
AUTHORS     Muzny,D.Marie., Metzker.M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,C.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,W., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,B., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,J., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Louiseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,R., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,D.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Oiarmpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,V., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
TITLE       Direct Submission
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 230556)
AUTHORS     Worley,K.C.
TITLE       Direct Submission
JOURNAL     Submitted (26-MAY-2005) Human Genome Sequencing Center, Department
```

```
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 230556)
AUTHORS     Cow Genome Sequencing Consortium.
TITLE       Direct Submission
JOURNAL     Submitted (01-JUL-2005) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Jun 29, 2005 this sequence version replaced gi:6730668.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table (a 'contig-scaffold'). Within each contig-scaffold,
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            --------------- Genome Center
            Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu/
            Contact: hgsc-help@bcm.tmc.edu
            --------------- Project Information
            Center project name: PGWN
            Center clone name: CH240-112A3
            --------------- Summary Statistics
            Assembly program: Atlas 3.0;
            Consensus quality: 223197 bases at least Q40
            Consensus quality: 225095 bases at least Q30
            Consensus quality: 225591 bases at least Q20
            Estimated insert size: 224127; sum-of-contigs estimation
            Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
            ----------------------------------------------------------------
          * NOTE: Estimated insert size may differ from sequence length
          *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
          * NOTE: This is a 'working draft' sequence. It currently
          *   consists of 24 contigs. The true order of the pieces
          *   is not known and their order in this sequence record is
          *   arbitrary. Gaps between the contigs are represented as
          *   runs of N, but the exact sizes of the gaps are unknown.
          *   This record will be updated with the finished sequence
          *   as soon as it is available and the accession number will
          *   be preserved.
          *      1      28520: contig of 28520 bp in length
          *  28520      28570: gap of 50 bp
          *  28571      41303: contig of 12733 bp in length
          *  41304      41353: gap of 50 bp
          *  41354      45169: contig of 3816 bp in length
          *  45170      45219: gap of 50 bp
          *  45220      75749: contig of 30530 bp in length
          *  75750      75799: gap of 50 bp
          *  75800      82009: contig of 6210 bp in length
          *  82010      82059: gap of 50 bp
          *  82060      85374: contig of 3315 bp in length
          *  85375      85586: gap of 212 bp
          *  85587     126194: contig of 40608 bp in length
          * 126195     126294: gap of unknown length
          * 126295     128873: contig of 2579 bp in length
          * 128874     128923: gap of 50 bp
          * 128924     129987: contig of 1064 bp in length
          * 129988     130087: gap of unknown length
          * 130088     137866: contig of 7779 bp in length
          * 137867     137916: gap of 50 bp
          * 137917     139075: contig of 1159 bp in length
          * 139076     139125: gap of 50 bp
          * 139126     150021: contig of 10896 bp in length
          * 150022     150456: gap of 435 bp
          * 150457     167878: contig of 17422 bp in length
          * 167879     167928: gap of 50 bp
```